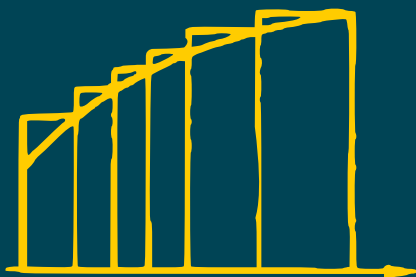


G. Chilov

Analyse Mathématique

Fonctions d'une Variable

1^{ère} et 2^e Parties



Editions Mir Moscou

G. CHILOV

ANALYSE MATHÉMATIQUE

**FONCTIONS
D'UNE VARIABLE**

TOMO I



Г. Е. ШИЛОВ

МАТЕМАТИЧЕСКИЙ АНАЛИЗ
ФУНКЦИИ ОДНОГО ПЕРЕМЕННОГО

Части 1—2

ИЗДАТЕЛЬСТВО «НАУКА» МОСКВА

G. CHILOV

Analyse mathématique

FONCTIONS

D'UNE

VARIABLE

1^{re} et 2^e parties

Deuxième édition

Editions Mir · Moscou

Réimpression 1978

**Traduit du russe par
Vitali Kharine**

На французском языке

© Traduction française Editions Mlr 1973

Avant-propos

L'Analyse est un vaste domaine des mathématiques lié aux notions de fonction, de dérivée et d'intégrale. A l'heure actuelle, elle embrasse nombre de domaines plus restreints : équations différentielles (ordinaires et à dérivées partielles), équations intégrales, fonctions de variable complexe, géométrie différentielle, calcul des variations et d'autres. Mais tandis que le contenu de l'analyse mathématique peut être considéré comme formé, les opinions sur sa structure subissent d'importants changements. Dans le cours classique des années 1920 de E. Goursat, toute l'analyse est présentée comme sur une immense plaine, à un même niveau d'abstraction ; dans les livres d'aujourd'hui, on prête une grande attention à la mise en relief des « étages » d'abstraction, c'est-à-dire des « structures » (Bourbaki) différentes qui caractérisent les bases mathématico-logiques des constructions initiales. Le retour aux bases éclaire l'essentiel du problème en libérant le mathématicien de la nécessité d'étudier les traits individuels de l'objet, ce qui, à son tour, rend possible l'étude de nouveaux objets offrant une autre individualité mais la même construction de fond.

Ainsi, on connaissait la démonstration de l'existence et de l'unicité de la solution d'une équation différentielle due à Picard et basée sur la méthode des approximations successives de la fonction cherchée sur un intervalle par d'autres fonctions obtenues selon des règles données. Ensuite, on a formulé (Banach et d'autres savants) la « méthode du point fixe » que l'on a appliquée à la démonstration du même théorème. Cette méthode a découvert la partie essentielle de la démonstration de Picard : l'existence d'un opérateur contractant dans un espace métrique. Et tout l'entourage, fonctions numériques sur un intervalle, équation différentielle, s'est avéré peu important. Finalement, la « méthode du point fixe » a non seulement rendu la démonstration du théorème de Picard plus claire, plus « géométrique », mais a aussi donné une possibilité de démontrer, en développant son idée, une quantité de théorèmes d'existence où il ne s'agit même pas de fonctions sur un intervalle ni d'équations

différentielles. On peut en dire de même de la géométrie de l'espace d'Hilbert, du calcul des fonctionnelles différentiables, etc.

Dans le présent livre, nous exposons les concepts fondamentaux de l'analyse mathématique relativement aux fonctions d'une variable. Mais nous attribuons au terme « une variable » un sens élargi. C'est que, pour certaines notions fondamentales de l'analyse, telles que limite et continuité, la distinction entre les cas classiques d'une et de plusieurs variables n'est pas si importante; pour les chapitres correspondants, nous préférons donc un niveau optimal de généralité en sous-entendant par « une variable » un point d'un espace métrique, par exemple. Cependant, lorsqu'on arrive à la dérivation et à l'intégration, la différence entre les cas classiques mentionnés devient sensible, et nous nous bornons alors aux fonctions « vraiment » d'une variable, d'abord réelle, puis complexe. Néanmoins, ce n'est qu'au début que les valeurs de ces fonctions sont numériques; par la suite elles deviennent vectorielles, même appartenant à un espace normé, ce qui offre un vaste champ d'applications. Les fonctions analytiques font dans notre exposé partie intégrante de l'analyse. Nous ne touchons pas, dans ce livre, au grand domaine du calcul différentiel et intégral des fonctions de plusieurs variables qui demanderait encore un volume au moins.

Nous n'avons pas inclus dans le livre l'intégrale de Lebesgue, parce que, dans les problèmes analytiques que l'on traite ici, on ne rencontre que les fonctions continues (ou possédant un nombre fini de points de discontinuité) pour l'intégration desquelles l'intégrale de Riemann est bien suffisante. En ce qui concerne les problèmes plus compliqués de l'analyse, par exemple la théorie des équations intégrales, le rôle décisif de l'intégrale de Lebesgue est incontestable. Mais un exposé de la théorie de l'intégrale de Lebesgue dans le présent livre risquerait de surcharger l'attention du lecteur par les nuances spécifiques de la théorie des fonctions de variable réelle et de celle de la mesure.

Bien que, formellement, le livre n'exige que les connaissances du cours secondaire, il serait bien si le lecteur s'était familiarisé avec la construction de graphiques, l'intégration, la dérivation et leurs applications géométriques les plus simples. Le présent ouvrage n'est pas un manuel d'analyse mathématique: il est plutôt destiné à susciter la réflexion, à faire comparer les aspects différents

de la théorie. Certains endroits peuvent être utilisés dans un cours ou un séminaire de type approfondi. Les exercices poursuivent le même but; aucun d'entre eux ne vise à travailler la technique (les recueils usités contiennent de tels exercices en quantité suffisante), mais tous illustrent et développent la théorie générale exposée.

Cet ouvrage est composé de trois parties. La première (« Introduction à l'analyse ») et la deuxième (« Calcul différentiel et intégral ») sont à la disposition du lecteur; la troisième partie « Chapitres choisis de l'analyse moderne » paraîtra en français fin 1973.

Dans la première partie, l'exposé systématique commence par la théorie des nombres réels (chapitre 1). Nous sous-entendons par nombres réels un ensemble d'objets satisfaisant à certains axiomes. Il existe d'autres constructions de la théorie des nombres réels où l'on démontre les propositions que nous choisissons pour axiomes en partant (dans un exposé rigoureux, par exemple dans le cours connu de Landau) des axiomes des nombres naturels et de ceux de la théorie des ensembles. Dans les constructions des deux types, il manque un élément essentiel, à savoir la démonstration de la non-contradiction des axiomes. Il semble que ce défaut soit propre à toutes les constructions modernes de la théorie des nombres réels. Le problème est loin d'être technique, il heurte les bases mêmes de la pensée mathématique. Dans ce contexte, le choix du point de départ pour le schéma général de l'analyse devient inessential, et nous le voulons le plus proche possible des constructions analytiques proprement dites.

Dans le chapitre 2, après une brève incursion dans la théorie des ensembles, nous introduisons les notions de structure mathématique et d'isomorphisme. En tant qu'illustration, nous établissons l'unicité (à un isomorphisme près) de la structure des nombres réels. Nous introduisons également les structures de l'espace n -dimensionnel et du corps des nombres complexes.

Le chapitre 3 est consacré à la théorie des espaces métriques. Dans le chapitre 4, nous développons la théorie générale des limites. Cette théorie a pour base, d'une part, un ensemble E sur lequel on choisit une direction (c.-à-d. une famille ordonnée d'ensembles d'intersection vide qui représente une formation moins générale que le filtre de Cartan mais bien suffisante pour l'analyse) et, d'autre part, une fonction définie sur l'ensemble E à valeurs dans un espace

métrique. Ce schéma embrasse tous les types de limites considérés dans l'analyse, à partir de la limite d'une suite numérique jusqu'à la dérivée et l'intégrale.

Dans le chapitre 5, après les théorèmes sur la continuité des fonctions numériques sur l'axe réel, nous introduisons, à l'aide d'équations fonctionnelles, le logarithme (dont l'inversion conduit à l'exponentielle) et les fonctions trigonométriques. L'algèbre et la topologie des nombres complexes ainsi que le théorème d'existence d'une racine d'un polynôme à coefficients complexes sont considérés en tant qu'applications.

Dans le chapitre 6, nous envisageons la théorie des séries (numériques, de puissances, fonctionnelles).

La deuxième partie du livre s'ouvre par le chapitre 7 sur la dérivée. Les chapitres 7 et 8 contiennent le calcul différentiel proprement dit. La formule et la série de Taylor amènent à l'extension naturelle de l'analyse réelle sur le domaine complexe.

Dans le chapitre 9, à côté de la théorie générale de l'intégrale de Riemann, on en trouve quelques applications. Pour le développement ultérieur de l'analyse, la technique des fonctions analytiques devient indispensable, et nous l'exposons dans le chapitre 10. Les fonctions analytiques s'avèrent, en particulier, très utiles pour le calcul des intégrales impropres auxquelles on consacre le chapitre suivant.

Le système de numérotation devient clair on considérant l'exemple suivant : le symbole 10.37. b veut dire « chapitre 10, paragraphe 3, numéro 7, point b ». D'une manière analogue, la formule 10.37 (4) est la quatrième formule du numéro 10.37. Dans les limites d'un numéro, les formules sont numérotées dans l'ordre naturel. Les figures et les exercices sont numérotés dans les limites d'un chapitre.

L'auteur

L'analyse mathématique est une symphonie cohérente de l'infini.

D. Hilbert

**PREMIÈRE
PARTIE**

**Introduction
à l'analyse**

Nombres réels

Nous concluons donc qu'il n'y a aucuns nombres absurds, irrationels, irreguliers, inexplicables ou sourds; mais qu'il y a en eux telle excellence, et concordance, que nous avons matiere de modifier nuit et jour en leur admirable perfection.

Simon Stevin

§ 1.1. Premières notions sur les ensembles

1.11. Lorsqu'on considère plusieurs objets quelconques (« éléments »), on emploie les termes tels que « collection », « totalité », « ensemble ». On parle, par exemple, de l'ensemble des étudiants d'une salle d'étude, de l'ensemble des grains de sable d'une plage, de l'ensemble des sommets d'un polygone ou de celui de ses côtés. Chacun de ces ensembles est composé d'un certain nombre d'éléments (que l'on peut estimer bien qu'il soit peut-être impossible de l'établir exactement *)). De tels ensembles sont dits *finis*.

En mathématiques, on a souvent affaire aux ensembles dont les objets ne sont pas en nombre fini; l'ensemble des nombres naturels 1, 2, 3, . . . , ainsi que celui de tous les points d'un segment en sont les exemples les plus simples **). De tels ensembles sont dits *infinis*.

De même nous classons parmi les ensembles *ensemble vide* qui est un ensemble dénué d'éléments.

D'habitude, nous désignons un ensemble par l'une des lettres majuscules A , B , C , . . . et ses éléments par des lettres minuscules. La notation $a \in A$ (ou bien $A \ni a$) signifie que a est élément de l'ensemble A ; la notation $a \notin A$ ou $a \notin A$ ou encore $A \not\ni a$ veut dire que a n'est pas élément de A . La notation $A \subset B$ (ou $B \supset A$) signifie que tout élément de A est élément de B ; dans ce cas l'en-

*) « Il est des personnes, ô roi Gélon, qui pensent que le nombre des grains de sable est infini... Quant à moi, je vais faire voir par des démonstrations géométriques auxquelles tu ne pourras refuser ton assentiment, que parmi les nombres... il y en a qui excèdent le nombre des grains d'un volume de sable égal non seulement à la grandeur de la terre, mais encore à celui de l'univers entier » (Archimède, *L'arénaire*).

**) Les définitions rigoureuses des êtres considérées dans § 1.1 en tant qu'exemples seront données plus bas. (Ici ils n'ont qu'une valeur illustrative.)

semble A est appelé *sous-ensemble* de l'ensemble B . Le plus vaste des sous-ensembles de B est évidemment l'ensemble B , le plus étroit l'ensemble vide. N'importe quel autre sous-ensemble de B contient forcément des éléments de B , mais pas tous. Chacun de ces sous-ensembles est appelé *sous-ensemble propre*. Les symboles $\in, \ni, \subset, \supset$ sont appelés *symboles d'inclusion*. Si les inclusions $A \subset B$, $B \subset A$ ont lieu, cela veut dire que tout élément de l'ensemble A est élément de l'ensemble B et, réciproquement, tout élément de B est élément de A ; ainsi, les ensembles A et B se composent de mêmes éléments et sont donc confondus. Ce fait est noté $A = B$. La notation analogue pour les éléments $a = b$ veut dire tout simplement que a et b est un même élément.

Il existe des méthodes différentes de définir un ensemble. La plus simple est d'énumérer tous ses éléments, par exemple: $A = \{1, 2, \dots, 15\}$. Un autre procédé souvent utilisé consiste à énoncer les propriétés caractéristiques des éléments de l'ensemble, par exemple: $A = \{x : x^2 - 1 < 0\}$ est l'ensemble de tous les x qui vérifient l'inéquation suivant le deux-points.

1.12. Considérons deux opérations simples que l'on peut effectuer sur les ensembles: *réunion* et *intersection*.

Définissons d'abord la réunion. Soient des ensembles $A, B, C, \dots$. L'ensemble de tous les éléments dont chacun appartient à l'un au moins des ensembles $A, B, C, \dots$ s'appelle *réunion* des ensembles $A, B, C, \dots$.

Ainsi, la réunion de l'ensemble $A = \{6, 7, 8, \dots\}$ (de tous les nombres naturels supérieurs à 5) et de l'ensemble $B = \{3, 6, 9, \dots\}$ (de tous les nombres naturels multiples de 3) est l'ensemble

$$S = \{3, 6, 7, 8, 9, 10, \dots\}$$

(de tous les nombres naturels sauf 1, 2, 4 et 5).

Introduisons maintenant l'opération d'intersection. On appelle *intersection* des ensembles $A, B, C, \dots$ l'ensemble des éléments appartenant à *chacun* des ensembles donnés.

Ainsi, dans l'exemple précédent, l'intersection des ensembles

$$A = \{6, 7, 8, 9, 10, \dots\},$$

$$B = \{3, 6, 9, 12, \dots\}$$

est l'ensemble

$$D = \{6, 9, 12, \dots\}.$$

Il se peut que les ensembles $A, B, C, \dots$ n'aient aucun élément commun. Alors ils ont l'ensemble vide pour intersection; on dit dans ce cas que les ensembles $A, B, C, \dots$ sont *disjoints*. Par exemple, les trois ensembles numériques

$$A = \{1, 2\}, \quad B = \{2, 3\}, \quad C = \{1, 3\}$$

sont disjoints (bien que n'importe quels deux d'entre eux possèdent des éléments communs).

On peut considérer la réunion et l'intersection des ensembles qui sont en nombre fini ou non. Ainsi, on peut construire la réunion des ensembles dont chacun est formé de tous les points d'une droite passant par un point donné O du plan. Cette réunion est évidemment, l'ensemble de tous les points du plan. L'intersection des ensembles mentionnés est l'ensemble formé d'un seul point O .

Pour la réunion d'ensembles on utilise les symboles $\sum$ et $\cup$, de sorte que, par exemple, la notation

$$S = \sum_{v=1}^{\infty} A_v \quad \text{ou} \quad S = \bigcup_{v=1}^{\infty} A_v$$

désigne la réunion des ensembles $A_1, A_2, \dots, A_v, \dots$

Pour l'intersection d'ensembles on utilise les symboles $[]$ et $\cap$, de sorte que, par exemple, la notation

$$D = \bigcap_{v=1}^{\infty} A_v \quad \text{ou} \quad D = \bigcap_{v=1}^{\infty} A_v$$

désigne l'intersection des ensembles $A_1, A_2, \dots, A_v, \dots$

§ 1.2. Axiomes des nombres réels

La définition donnée ci-dessous part des propriétés élémentaires des nombres dont nous prenons connaissance par l'expérience quotidienne et dans l'école secondaire *). Au lieu de définir un nombre réel isolé, nous définissons tout l'ensemble des nombres réels comme ensemble muni de certaines relations et opérations.

Les propriétés des relations et des opérations sont données par un système d'axiomes.

Les axiomes sont partagés en quatre groupes : les axiomes d'addition, les axiomes de multiplication, les axiomes d'ordre et le quatrième groupe qui contient un seul axiome, celui de borne supérieure.

*) Le rôle de l'expérience quotidienne n'est pas tout aussi déterminant pour l'axiome de borne supérieure (1.24). Or, il en est de même du postulat d'Euclide (sur l'existence d'une parallèle unique) qui est à la base de la géométrie. L'expérience ne nous fournit pas des axiomes mathématiques d'une façon catégorique; entre l'expérience et la science, il y a une étape de formation d'axiomes qui, dans le cadre d'une même expérience, peuvent varier d'une façon radicale. Tout comme, parallèlement à la géométrie euclidienne, il existe une géométrie non euclidienne (celle de Lobatchevski) avec l'axiome d'existence de différentes droites passant par un point donné et parallèles à une droite donnée, il existe, à côté de la théorie des nombres réels que nous exposons, d'autres théories dans lesquelles un ensemble borné ne possède pas forcément une borne supérieure [6].

Définition. Un ensemble d'éléments $x, y, z, \dots$ est appelé *ensemble R des nombres réels* si les propriétés suivantes ont lieu :

1.21. Opération d'addition : à deux objets quelconques x, y il correspond un objet z appelé *somme* de x et y et noté $x + y$, de sorte que les conditions suivantes sont satisfaites :

- a. $x + y = y + x$ quels que soient x et y de R .
- b. $(x + y) + z = x + (y + z)$ quels que soient x, y, z de R ; c'est pourquoi l'expression $x + y + z$ est univoque.
- c. Il existe dans R un élément désigné par 0 (*zéro*) et tel que $x + 0 = x$ pour tout $x \in R$.
- d. Pour tout $x \in R$, il existe un élément y dit *opposé* de x et tel que $x + y = 0$.

1.22. Opération de multiplication : à deux objets quelconques x, y il correspond un objet $u \in R$ appelé *produit* de x et y et noté $x \cdot y$ (ou xy), de sorte que les conditions suivantes sont remplies :

- a. $xy = yx$ quels que soient x et y de R .
- b. $(xy)z = x(yz)$ quels que soient x, y, z de R ; c'est pourquoi l'expression xyz est univoque.
- c. Il existe dans R un élément différent de 0 , désigné par 1 (*unité*) et tel que $x \cdot 1 = x$ pour tout $x \in R$.
- d. Pour tout $x \neq 0$, il existe dans R un élément u , dit *inverse* de x , tel que $u \cdot x = 1$.
- e. Quels que soient x, y, z de R , l'égalité

$$x(y + z) = xy + xz$$

est valable.

Le dernier axiome lie l'opération de multiplication à celle d'addition introduite plus haut (1.21).

L'ensemble des objets $x, y, \dots$ satisfaisant aux axiomes 1.21-1.22 est appelé *corps numérique* ou *corps* tout court.

1.23. Relation d'ordre : Deux éléments quelconques x, y de R satisfont à l'une (ou aux deux à la fois) des relations $x \leq y$ (x est inférieur ou égal à y) et $y \leq x$ avec les propriétés suivantes :

- a. $x \leq x$ pour tout x ; $x \leq y, y \leq x$ impliquent $x = y$.
- b. $x \leq y, y \leq z$ impliquent $x \leq z$.
- c. $x \leq y$ implique $x + z \leq y + z$ pour tout z de R .
- d. $0 \leq x, 0 \leq y$ impliquent $0 \leq xy$.

La relation $x \leq y$ est également notée $y \geq x$ (y est supérieur ou égal à x). La relation $x \leq y$ pour $x \neq y$ est notée $x < y$ (x est inférieur à y) ou $y > x$ (y est supérieur à x).

1.24. Un ensemble $A \subset R$ est dit *majoré* s'il existe un élément $z \in R$ tel que $x \leq z$ pour tout $x \in A$; cette relation est notée $A \leq z$.

Tout nombre z possédant la propriété mentionnée par rapport à un ensemble A s'appelle *majorant* de A . Un majorant z_0 de l'ensemble A est appelé *borne supérieure* de A si tout autre majorant z de A est supérieur ou égal à z_0 . La borne supérieure d'un ensemble A est désignée par $\sup A$. A présent nous formulons le dernier axiome :

Axiome de borne supérieure. *Tout ensemble $A \subset R$ (non vide) majoré possède une borne supérieure.*

Nous allons établir des conséquences logiques des axiomes formulés plus haut ; la totalité de ces conséquences fournira toutes les propriétés du système des nombres réels que l'on utilise pour la construction de l'analyse mathématique.

§ 1.3. Conséquences des axiomes d'addition

1.31. *Il n'existe dans l'ensemble R qu'un seul zéro.*

En effet, supposons qu'il y en ait deux : 0_1 et 0_2 . Alors, en appliquant les axiomes 1.21.a et c, on obtient

$$0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2.$$

1.32. *Tout élément x de l'ensemble R n'admet qu'un seul élément opposé.*

Supposons qu'un x possède deux éléments opposés y_1 et y_2 , de sorte que $x + y_1 = x + y_2 = 0$. Alors, d'après les axiomes 1.21.a-c nous avons

$$\begin{aligned} y_2 &= 0 + y_2 = (x + y_1) + y_2 = x + (y_1 + y_2) = \\ &= x + (y_2 + y_1) = (x + y_2) + y_1 = 0 + y_1 = y_1. \end{aligned}$$

L'élément opposé d'un élément x se note $-x$. La somme $x + (-y)$ est aussi notée $x - y$ et s'appelle *différence* de x et y . L'opposé d'une somme est la somme des opposés de ses termes : en effet, $-x - y + (x + y) = -x - y + x + y = -x + x - y + y = 0 + 0 = 0$.

1.33. *L'équation*

$$a + x = b \tag{1}$$

possède dans R la solution unique $b - a$.

En effet, si l'on ajoute le nombre $-a$ à chaque membre de l'équation (1), on trouve en appliquant les axiomes 1.21.a-c :

$$a + x - a = a - a + x = 0 + x = x = b - a,$$

de sorte que si une solution existe, elle est égale à $b - a$. Or $b - a$ est bien une solution car

$$a + (b - a) = a + b + (-a) = a + (-a) + b = 0 + b = b.$$

§ 1.4. Conséquences des axiomes de multiplication

1.41.a. Il n'existe dans l'ensemble R qu'une seule unité.

Supposons qu'il y en ait deux: 1_1 et 1_2 . Alors, en appliquant l'axiome 1.22.a, nous obtenons

$$1_1 = 1_1 \cdot 1_2 = 1_2.$$

b. Tout élément $x \neq 0$ de l'ensemble R n'admet qu'un seul élément inverse.

Supposons qu'un x a deux inverses z_1 et z_2 , de sorte que $xz_1 = 1$, $xz_2 = 1$. Alors, d'après les axiomes 1.22.a-c, on a

$$z_2 = 1 \cdot z_2 = (xz_1) z_2 = x (z_1 z_2) = x (z_2 z_1) = (xz_2) z_1 = 1 \cdot z_1 = z_1.$$

1.42. L'élément inverse d'un élément x se note $\frac{1}{x}$. L'élément $\frac{1}{xy}$ inverse d'un produit xy est égal au produit des éléments inverses de x et y : en effet,

$$\frac{1}{x} \cdot \frac{1}{y} \cdot xy = \frac{1}{x} \cdot x \cdot \frac{1}{y} \cdot y = 1 \cdot 1 = 1.$$

Le produit $x \cdot \frac{1}{z}$ est aussi noté $\frac{x}{z}$ et appelé *quotient* de x par z .

1.43. Définition. Les nombres $1, 2 = 1 + 1, 3 = 2 + 1, \dots, n = (n-1) + 1, \dots$ sont dits *naturels*. L'ensemble des nombres naturels peut donc être défini comme le plus petit ensemble contenant le nombre 1 et, avec tout nombre n , le nombre $n + 1$.

Dans bien des problèmes on demande de prouver qu'un ensemble numérique A (par exemple, l'ensemble des nombres naturels n qui satisfont à une proposition T_n dépendant de n) contient *tous* les nombres naturels. Le raisonnement par récurrence que l'on applique à de tels problèmes consiste à vérifier les conditions suivantes:

- 1) A contient le nombre 1;
- 2) si A contient un nombre naturel n , il contient aussi le nombre $n + 1$.

Ceci dit, il est clair que l'ensemble A contient tous les nombres naturels.

Nous voyons que le bien-fondé de cette méthode de démonstration découle de la définition même des nombres naturels.

1.44.a. Les nombres naturels, leurs opposés et le zéro sont appelés nombres *entiers*.

b. Soit m un entier; alors l'entier $2m$ est dit *pair*, $2m + 1$ *impair*.

c. Les quotients $\frac{m}{n}$, où m, n sont entiers et $n \neq 0$, s'appellent *nombre rationnels*.

d. Tous les autres nombres réels sont dits *irrationnels*.

1.45. L'équation

$$ax = b \quad (a \neq 0) \quad (1)$$

possède dans R la solution unique $\frac{b}{a}$.

En effet, en multipliant chaque membre de l'équation (1) par $\frac{1}{a}$ on trouve

$$\frac{1}{a}(ax) = \left(\frac{1}{a} \cdot a\right)x = 1 \cdot x = x = \frac{b}{a}.$$

de sorte que si une solution existe, elle est égale à $\frac{b}{a}$. Or $\frac{b}{a}$ est bien une solution car

$$a \cdot \frac{b}{a} = \left(a \cdot \frac{1}{a}\right)b = 1 \cdot b = b.$$

1.46. On pose par définition

$$x^1 = x, \quad x^n = \underbrace{x \dots x}_{n \text{ fois}} \quad \text{pour } n = 2, 3 \dots$$

Evidemment, $x^n \cdot x^m = x^{n+m}$ et $(x^n)^m = x^{nm}$ quels que soient n et m naturels. Le nombre n dans l'expression x^n est appelé *exposant de la puissance*. Étendons la notion d'exposant sur tous les nombres entiers. Pour tout $x \neq 0$, posons

$$x^0 = 1, \quad x^{-n} = \frac{1}{x^n}$$

Vérifions que les formules

$$x^n \cdot x^m = x^{n+m}, \quad (x^n)^m = x^{nm} \quad (1)$$

restent valables quels que soient n et m entiers. Soit $n > 0$, $m = -p < 0$, où $p \leq n$; alors

$$x^n \cdot x^m = x^n \cdot \frac{1}{x^p} = x^{n-p} \cdot x^p \cdot \frac{1}{x^p} = x^{n-p} = x^{n+m}.$$

Si $p > n$, alors, d'après 1.42, on a

$$x^n \cdot x^m = x^n \cdot \frac{1}{x^p} = x^n \cdot \frac{1}{x^p x^{p-n}} = x^n \cdot \frac{1}{x^p} \cdot \frac{1}{x^{p-n}} = \frac{1}{x^{p-n}} = x^{n+m}.$$

Et si $n = -p < 0$, $m = -q < 0$, alors, toujours d'après 1.42, on a

$$x^n \cdot x^m = \frac{1}{x^p} \cdot \frac{1}{x^q} = \frac{1}{x^{p+q}} = x^{m+n}.$$

On vérifie de façon analogue la deuxième formule (1).

1.47.a. Pour tout $x \in R$ a lieu l'égalité :

$$0 \cdot x = 0.$$

En effet,

$$0 \cdot x + 1 \cdot x = (0 + 1) x = 1 \cdot x = x,$$

$$0 \cdot x + 1 \cdot x = 0 \cdot x + x,$$

d'où

$$x = 0 \cdot x + x;$$

en vertu de 1.33, $0 \cdot x = x - x = 0$.

Il en découle que 0 n'a pas d'inverse puisque l'égalité $0 \cdot x = 1$ est impossible.

b. D'autre part, il résulte de $xy = 0$ et $x \neq 0$ que $y = \left(\frac{1}{x} \cdot x\right) y = \frac{1}{x} (xy) = \frac{1}{x} \cdot 0 = 0$. Donc, si un produit est nul, l'un au moins des facteurs est nul.

1.48. Quels que soient $u \neq 0$, $v \neq 0$, on a :

$$\frac{x}{u} + \frac{y}{v} = \frac{xv + yu}{uv}.$$

En effet,

$$\frac{xv + yu}{uv} = \frac{1}{uv} (xv + yu) = \frac{1}{uv} \cdot xv + \frac{1}{uv} \cdot yu = \frac{x}{u} + \frac{y}{v}.$$

1.49. Pour tout $x \in R$, on a :

$$-x = (-1)x.$$

Notons que les deux membres de l'égalité sont définis indépendamment l'un de l'autre, donc l'égalité est à démontrer. Nous avons selon 1.47.a :

$$(-1)x + x = [(-1) + 1] \cdot x = 0 \cdot x = 0,$$

d'où le résultat demandé.

Grâce aux conséquences des §§ 1.3-1.4 toutes les identités de l'algèbre élémentaire des nombres réels (binôme de Newton, formules de sommation des progressions, propriétés des déterminants, etc.) se trouvent établies.

§ 1.5. Conséquences des axiomes d'ordre

1.51. Relations entre l'ordre et l'opération d'addition.

a. Si $x \leq y$, $y \leq z$ et $x = z$, alors $x = y = z$.

En effet, $y \leq z = x$, donc $y \leq x$, d'où, d'après l'axiome 1.23.a, $y = x$.

Il résulte directement de a :

b. $x < y$, $y \leq z$ impliquent $x < z$. D'une façon analogue, $x \leq y$, $y < z$ impliquent $x < z$.

c. Les relations $x \leq y$, $0 \leq y - x$, $-y \leq -x$, $x - y \leq 0$ sont équivalentes.

En effet, en ajoutant $-x$ aux deux membres de la première inégalité et en appliquant 1.23.c, on obtient la deuxième; en ajoutant $-y$ aux deux membres de la deuxième inégalité, on obtient la troisième; en ajoutant x aux deux membres de la troisième inégalité, on obtient la quatrième et en ajoutant y aux deux membres de la quatrième, on retrouve la première inégalité.

d. $x < y$ implique $x + z < y + z$ pour tout $z \in R$. En effet, $x < y$ entraîne évidemment $x \leq y$ et $x + z \leq y + z$. Or, si l'égalité $x + z = y + z$ avait lieu, alors, en ajoutant $-z$ à ses deux membres, on aurait $x = y$, ce qui contredit l'hypothèse. Donc $x + z < y + z$.

e. Si $x_1 \leq y_1, \dots, x_n \leq y_n$, alors $x_1 + \dots + x_n \leq y_1 + \dots + y_n$; de plus, si l'inégalité $x_j < y_j$ est juste pour au moins deux x_j, y_j quelconques, alors $x_1 + \dots + x_n < y_1 + \dots + y_n$.

En effet, on a d'après l'axiome 1.23.c:

$$x_1 + \dots + x_n \leq y_1 + x_2 + \dots + x_n \leq y_1 + y_2 + \dots + x_n \leq \dots \leq y_1 + \dots + y_n;$$

de plus, si l'on a $x_j < y_j$ pour au moins deux x_j, y_j quelconques, alors, en vertu de d, le symbole $<$ apparaît à l'endroit correspondant de la chaîne et sera conservé dans la suite, toujours d'après d. Ainsi, on peut additionner les inégalités d'un même sens. En particulier, $x_1 \leq 0, \dots, x_n \leq 0$ implique $s = x_1 + \dots + x_n \leq 0$; de plus, si l'on a $x_j < 0$ pour au moins un j , alors on a aussi $s < 0$. Le fait analogue a lieu en remplaçant $\leq$ par $\geq$ et $<$ par $>$.

f. Les relations $x < y$, $0 < y - x$, $-y < -x$, $x - y < 0$ sont équivalentes.

Cela se déduit de e tout comme c a été déduit de l'axiome 1.23.c.

1.52. Définition. Si $x \geq 0$ ($x > 0$), le nombre x est dit *non négatif* (*positif*); si $x \leq 0$ ($x < 0$), le nombre x est dit *non positif* (*négatif*). Le nombre 0 est non positif et non négatif à la fois.

1.53. Définition. Soient deux nombres réels x et y et, par exemple, $x \leq y$. Alors x est dit *minimal* des nombres x et y et se note $x = \min \{x, y\}$; y est dit *maximal* des nombres x et y et se note $y = \max \{x, y\}$. On peut définir par récurrence $\min \{x_1, \dots, x_n\}$ et $\max \{x_1, \dots, x_n\}$ pour tout ensemble (fini) des nombres $x_1, \dots, x_n$ (par exemple, $\max \{x_1, \dots, x_n\} = \max \{\max \{x_1, \dots, x_{n-1}\}, x_n\}$).

Le nombre $|x| = \max \{x, -x\}$ est appelé *module* ou *valeur absolue* du nombre x . De la sorte, $|x| = x$ pour $x \geq 0$ et $|x| = -x$ pour $x \leq 0$; pour tout x , le nombre $|x|$ est non négatif et $|-x| = |x|$.

1.54.a. Pour $a > 0$ l'inégalité $|x| \leq a$ est équivalente à deux inégalités $x \leq a$, $-x \leq a$ ou bien à $-a \leq x \leq a$.

b. Quels que soient deux nombres réels x , y , on a :

$$|x + y| \leq |x| + |y|. \quad (1)$$

En effet, si x , y sont les deux non négatifs ou non positifs, alors l'inégalité est juste selon la définition du module. Et si, par exemple, $x \geq 0$ et $y \leq 0$, alors :

$$x + y \leq x \leq x + |y| = |x| + |y|,$$

$$-x - y \leq -y = |y| \leq |x| + |y|.$$

de sorte que $|x + y| = \max \{x + y, -x - y\} \leq |x| + |y|$, ce qu'il fallait démontrer.

c. Il résulte de (1) par récurrence que

$$|x_1 + \dots + x_n| \leq |x_1| + \dots + |x_n|.$$

1.55. Relations entre l'ordre et l'opération de multiplication.

a. Si $x > 0$, $y > 0$, alors $xy > 0$.

b. Si $x \leq y$, alors, pour tout $z > 0$, on a :

$$xz \leq yz.$$

La proposition a résulte de l'axiome 1.23.d compte tenu de 1.47.b. La proposition b découle de l'inégalité

$$yz - xz = (y - x)z \geq 0$$

en vertu de 1.23.d.

c. En appliquant 1.47.b, on peut remplacer $\leq$ par $<$ partout dans b.

d. En particulier, on a $x^2 > x$ pour $x > 1$ et $x^2 < x$ pour $0 < x < 1$.

e. Si $x \leq y$, $0 < z \leq u$, alors $xz \leq yz \leq yu$. Aussi peut-on, sous les conditions formulées, multiplier les inégalités.

f. En particulier, pour $0 < x < y$ on a toujours $x^2 < y^2$, ...
..., $x^n < y^n$.

g. Si $x \leq 0$, $y \geq 0$, alors $xy \leq 0$; si $x \leq 0$, $y \leq 0$, alors $xy \geq 0$.

En effet, avec la première hypothèse on a $-x \geq 0$, et, d'après l'axiome 1.23.d et la propriété 1.49, on a $(-x)y = -1 \cdot x \cdot y = -(xy) \geq 0$, d'où $xy \leq 0$. Avec la seconde hypothèse on a $-y \geq 0$ et, en appliquant le résultat précédent, on obtient $-xy \leq 0$, $xy \geq 0$. Dans tous ces cas, on peut remplacer $\leq$ par $<$.

h. En particulier, pour tout $x \neq 0$, on a $x^2 = x \cdot x > 0$. D'où $1 = 1 \cdot 1 > 0$; ensuite, d'après 1.51.d, $2 = 1 + 1 > 1 + 0 = 1$, $3 = 2 + 1 > 2$, etc.

i. Quels que soient x , y , on a l'égalité $|x \cdot y| = |x| \cdot |y|$.

1.56. Si $x > 0$, alors $\frac{1}{x} > 0$; de $0 < x < y$ il résulte $0 < \frac{1}{y} < \frac{1}{x}$.

La première proposition découle de $x \cdot \frac{1}{x} = 1 > 0$ et de 1.55.g. En multipliant l'inégalité $0 < x < y$ par $\frac{1}{xy}$, nous obtenons la seconde proposition. En particulier, tous les nombres rationnels $\frac{p}{q}$, p et q étant des nombres naturels, sont positifs.

1.57. Le principe suivant est souvent appliqué dans des démonstrations :

Si un nombre z est non négatif et inférieur à tout nombre positif, alors $z = 0$.

En effet, si $z > 0$, alors nous avons forcément par hypothèse $z < z$, ce qui est impossible (cf. 1.23).

§ 1.6. Conséquences de l'axiome de borne supérieure

L'ensemble R de tous les nombres réels sera également appelé *droite numérique* et les nombres réels mêmes *points de cette droite*.

1.61. Dans 1.24 on a donné la définition d'un ensemble majoré. Considérons à présent la notion d'ensemble minoré. Un ensemble $E \subset R$ est dit *minoré* s'il existe un élément $z \in R$ tel que $z \leq x$ pour tout $x \in E$; cette relation est notée $z \leq E$. Tout nombre z possédant la propriété signalée par rapport à un ensemble E s'appelle *minorant de l'ensemble E* . Si E est majoré, i.e. s'il existe un y tel que $E \leq y$, alors l'ensemble $-E$ (l'ensemble des nombres $-x$ pour $x \in E$) est minoré puisque $x \leq y$ implique $-x \geq -y$; de plus, $-y$ est un minorant de l'ensemble $-E$. Réciproquement, si E est minoré, alors $-E$ est majoré d'après les mêmes raisonnements, et si η est un minorant de l'ensemble E , alors $-\eta$ est un majorant de l'ensemble $-E$.

Un minorant y_0 d'un ensemble E minoré s'appelle *borne inférieure de l'ensemble E* si tout autre minorant de l'ensemble E est inférieur ou égal à y_0 . La borne inférieure d'un ensemble E est désignée par $\inf E$.

Théorème. *Tout ensemble E minoré possède la borne inférieure qui est égale à $-\sup(-E)$.*

En effet, l'ensemble $-E$ est majoré, et il existe, d'après l'axiome 1.24, le nombre $\xi = \sup(-E)$. Montrons que $-\xi = \inf E$. Pour tout $x \in E$ on a $-x \leq \xi$, d'où $-\xi \leq x$; donc $-\xi$ est bien un minorant de l'ensemble E . Soit η un autre minorant de l'ensemble E . Alors $-\eta$ est un majorant de l'ensemble $-E$ et, par définition, $-\eta \geq \sup(-E) = \xi$; d'où $\eta \leq -\xi$, ce qu'il fallait démontrer.

1.62.a. Si E et F sont majorés et $E \subset F$, alors $\sup E \leq \sup F$; si E et F sont minorés et $E \subset F$, alors $\inf E \geq \inf F$.

En effet, dans le premier cas $\sup F$ est un majorant pour F et à fortiori pour $E \subset F$; donc $\sup E \leq \sup F$. De même, dans le deuxième cas $\inf F$ est un minorant pour F et à fortiori pour $E \subset F$, donc $\inf F \leq \inf E$.

b. Si, pour n'importe quels $x \in E$ et $y \in F$, on a l'inégalité $x \leq y$, alors E est majoré, F minoré et $\sup E \leq \inf F$.

En effet, l'ensemble E est majoré par tout $y \in F$, donc $\sup E$ existe et $\sup E \leq y$ pour tout $y \in F$. Il en résulte que F est minoré par le nombre $\sup E$; donc $\sup E \leq \inf F$.

1.63. A présent on va démontrer l'existence et l'unicité de la racine n -ième de tout nombre positif.

Théorème. Pour tout $x > 0$ réel et pour tout $n > 0$ entier, il existe un et seulement un $y > 0$ réel tel que $y^n = x$.

Ce nombre est noté $\sqrt[n]{x}$ (racine n -ième de x).

Démonstration [21]. Considérons l'ensemble A de tous les z positifs et tels que $z^n \leq x$. Cet ensemble est majoré (par le nombre 1, si $x \leq 1$, et par le nombre x , si $x \geq 1$). Posons

$$y = \sup A. \quad (1)$$

Montrons que $y^n = x$. Soit $y^n < x$, $x - y^n = e$. Pour tout $h \leq 1$ positif, on a d'après la formule du binôme de Newton :

$$\begin{aligned} (y+h)^n &= y^n + ny^{n-1}h + \frac{n(n-1)}{1 \cdot 2} h^2 y^{n-2} + \dots = \\ &= y^n + h \left[ny^{n-1} + \frac{n(n-1)}{1 \cdot 2} h y^{n-2} + \dots \right] \leq \\ &\leq y^n + h \left[ny^{n-1} + \frac{n(n-1)}{1 \cdot 2} y^{n-2} + \dots \right] = \\ &= y^n + h [(1+y)^n - y^n]. \end{aligned}$$

On peut poser

$$h < \frac{e}{(1+y)^n - y^n},$$

après quoi on obtient $(y+h)^n \leq y^n + e = x$, ce qui contredit la définition (1). Donc $y^n \geq x$. Soit $y^n > x$, $y^n - x = e$. Pour tout $h \leq 1$ positif on a :

$$\begin{aligned} (y-h)^n &= y^n - ny^{n-1}h + \frac{n(n-1)}{1 \cdot 2} y^{n-2}h^2 - \dots = \\ &= y^n - h \left[ny^{n-1} - \frac{n(n-1)}{1 \cdot 2} y^{n-2}h + \dots \right] \geq \\ &\geq y^n - h \left[ny^{n-1} + \frac{n(n-1)}{1 \cdot 2} y^{n-2}h + \dots \right] \geq \end{aligned}$$

$$\begin{aligned} &\geq y^n - h \left[ny^{n-1} + \frac{n(n-1)}{1 \cdot 2} y^{n-2} + \dots \right] = \\ &= y^n - h[(1+y)^n - y^n]. \end{aligned}$$

On peut poser toujours

$$h < \frac{\varepsilon}{(1+y)^n - y^n},$$

et alors on obtient

$$(y-h)^n \geq y^n - \varepsilon = x,$$

ce qui contredit toujours la définition (1). Donc $y^n = x$, ce qu'il fallait démontrer.

L'unicité de la racine suit de l'inégalité $y_1^n < y_2^n$ pour $y_1 < y_2$ (1.55.f).

1.64. Quels que soient deux nombres positifs x et y , on a

$$\sqrt[n]{xy} = \sqrt[n]{x} \cdot \sqrt[n]{y}. \quad (1)$$

Soient $\xi = \sqrt[n]{x}$, $\eta = \sqrt[n]{y}$, $\tau = \sqrt[n]{xy}$. Nous avons $\xi^n = x$, $\eta^n = y$; alors $(\xi\eta)^n = \xi^n \eta^n = xy = \tau^n$. En vertu de l'unicité de la racine, on a :

$$\tau = \sqrt[n]{xy} = \xi\eta,$$

ce que nous cherchions.

On démontre de façon analogue que pour un x positif et pour deux entiers $m, n > 1$:

$$\sqrt[n]{\sqrt[m]{x}} = \sqrt[nm]{x}. \quad (2)$$

1.65. Etant donné $(-x)^n = (-1)^n x^n$, l'équation $y^n = x > 0$ possède, pour n pair, la solution négative $y_2 = -\sqrt[n]{x}$ à côté de la solution positive $y_1 = \sqrt[n]{x}$; et l'équation $y^n = x < 0$ n'a aucune solution réelle. Pour n impair, l'équation $y^n = x > 0$ possède, dans le domaine des nombres réels, la solution unique $y = \sqrt[n]{x}$. Dans ce cas l'équation $y^n = x < 0$ elle aussi possède la solution unique $y = -\sqrt[n]{|x|}$.

Les formules 1.64 (1), (2) permettent de construire d'une façon ordinaire toute l'algèbre élémentaire des expressions contenant des racines de nombres réels et, en particulier, les formules de résolution des équations du second degré ainsi que des équations plus compliquées, l'objet de l'algèbre élémentaire.

1.66. Un ensemble E majoré et minoré en même temps est dit *borné*. Tout ensemble borné E possède la borne supérieure $\sup E$ et la borne inférieure $\inf E$.

Les intervalles servent d'exemples d'ensembles bornés.

Pour $a < b$, l'ensemble de tous les nombres réels x vérifiant l'inégalité $a \leq x \leq b$ est appelé *intervalle fermé* et désigné par $[a, b]$. L'ensemble de tous les nombres réels x vérifiant l'inégalité $a < x < b$ s'appelle *intervalle ouvert* et se désigne par (a, b) . Un

intervalle fermé comprend ses extrémités, les extrémités d'un intervalle ouvert ne lui appartiennent pas. Néanmoins,

$$\sup [a, b] = \sup (a, b) = b, \quad \inf [a, b] = \inf (a, b) = a.$$

Il y a encore les intervalles « semi-ouverts »; ainsi, l'ensemble $\{x: a < x \leq b\} = (a, b]$ s'appelle *intervalle semi-ouvert à gauche*, et l'ensemble $\{x: a \leq x < b\} = [a, b)$ *intervalle semi-ouvert à droite*.

Parfois, pour unifier la terminologie, un point est aussi appelé intervalle et noté $a = [a, a] = \{x: a \leq x \leq a\}$.

§ 1.7. Principe d'Archimède et ses conséquences

1.71. Si x est un nombre réel et n un nombre entier, alors les nombres nx s'appellent *multiples entiers de x* .

Principe d'Archimède *). Si $x > 0$ et y est un nombre réel quelconque, alors il existe un multiple entier nx du nombre x tel que $(n-1)x \leq y$, $nx > y$.

Démonstration. Supposons que l'inégalité $px \leq y$ ait lieu pour tous les p entiers. Ceci veut dire que l'ensemble A de tous les nombres $\{px\}$ est majoré et a y pour majorant. D'après l'axiome 1.24, il existe la borne supérieure $\xi = \sup A$ de l'ensemble $\{px\}$. Le nombre $\xi - x < \xi$ n'est plus un majorant de l'ensemble A ; il existe donc un p tel que $px > \xi - x$. D'où $(p+1)x > \xi$ et ξ ne peut pas être un majorant de l'ensemble A . La contradiction obtenue démontre l'existence d'un nombre entier p pour lequel $px > y$. D'une façon analogue, il existe un nombre entier q pour lequel $qx < y$; il est évident que $q \leq p$. En examinant tour à tour les couples $(q, q+1)$, $(q+1, q+2)$, ..., $(p-1, p)$ trouvons-en un, par exemple $(n-1, n)$, pour lequel $(n-1)x \leq y$ et $nx > y$.

En particulier, si $x = 1$, nous obtenons que, pour tout $y \in \mathbb{R}$, il existe un n entier tel que $n-1 \leq y < n$. Le nombre $n-1$ s'appelle *partie entière* du nombre y et se désigne par $[y]$. Le nombre $y - [y]$ est appelé *partie fractionnaire* du nombre y et désigné par $\{y\}$. De la sorte, tout nombre y est la somme de sa partie entière et de sa partie fractionnaire: $y = [y] + \{y\}$.

1.72. En remplaçant partout dans 1.71 l'addition par la multiplication on obtient la version multiplicative suivante du principe d'Archimède:

Si $x > 1$, $y > 0$, alors il existe un exposant entier n tel que

$$x^{n-1} \leq y, \quad x^n > y.$$

*) Dans d'autres théories axiomatiques modernes des nombres réels, le principe d'Archimède, de même que celui de Cantor (§ 1.8), figure sur la liste d'axiomes; l'axiome de borne supérieure (1.24) devient alors théorème.

1.73. Si le nombre y , dans l'hypothèse du principe d'Archimède, est positif lui aussi, alors le nombre $n > \frac{y}{x}$ est positif. En multipliant la dernière inégalité par $\frac{x}{n}$ on arrive à la conclusion suivante :

Quels que soient $x > 0$ et $y > 0$, il existe un nombre naturel n tel que $\frac{y}{n} < x$.

Il en découle que, pour tout $y > 0$, on a :

$$\inf \left\{ \frac{y}{n}, n = 1, 2, \dots \right\} = 0. \quad (1)$$

En effet, l'ensemble dans (1) est formé de nombres positifs, donc sa borne inférieure est non négative. Or, on a vu qu'elle ne peut pas être positive; d'où (1).

1.74. C o n s é q u e n c e. *Chacun des systèmes d'intervalles semi-ouverts :*

$$(0, y) \supset \left(0, \frac{y}{2}\right] \supset \dots \supset \left(0, \frac{y}{n}\right] \supset \dots, \quad (1)$$

$$(a, a + y) \supset \left(a, a + \frac{y}{2}\right] \supset \dots \supset \left(a, a + \frac{y}{n}\right] \supset \dots, \quad (2)$$

$$(a - y, a) \supset \left[a - \frac{y}{2}, a\right) \supset \dots \supset \left[a - \frac{y}{n}, a\right) \supset \dots \quad (3)$$

a l'intersection vide ($y > 0$).

En effet, si les intervalles du système (2) avaient un point commun ξ , alors $\xi - a$ serait un point commun du système (1); et si les intervalles du système (3) avaient un point commun η , alors $a - \eta$ serait un point commun du système (1). Or, les intervalles du système (1) n'ont aucun point commun en vertu de 1.73, ce qui prouve la proposition.

1.75. T h é o r è m e. *Tout intervalle (a, b) contient un point rationnel.*

D é m o n s t r a t i o n. Soit $h = b - a > 0$ et soit n un entier supérieur à $\frac{1}{h}$ (il existe d'après le principe d'Archimède) de sorte que $\frac{1}{n} < h$. D'après le principe d'Archimède, il existe un m tel que $\frac{m}{n} \leq a < \frac{m+1}{n}$. On a $\frac{m+1}{n} - a \leq \frac{1}{n} < b - a$ de sorte que $\frac{m+1}{n} < b$. Ainsi, $a < \frac{m+1}{n} < b$, $\frac{m+1}{n} \in (a, b)$, ce qu'il fallait démontrer.

En réalité, il existe entre a et b tout un ensemble infini de nombres rationnels puisque, en appliquant le raisonnement ci-dessus

à l'intervalle $(\frac{m-1}{n}, b)$, on obtient un autre nombre rationnel $\frac{p}{q}$, $\frac{m-1}{n} < \frac{p}{q} < b$, et ce procédé peut continuer indéfiniment.

1.76. Pour un nombre réel donné ξ , désignons par N_ξ l'ensemble de tous les nombres rationnels $s \leq \xi$ et par P_ξ celui de tous les nombres rationnels $r \geq \xi$. L'ensemble N_ξ est majoré (par le nombre ξ), l'ensemble P_ξ minoré (par le nombre ξ).

T h é o r è m e. $\sup N_\xi = \xi = \inf P_\xi$.

D é m o n s t r a t i o n. Soit $\sup N_\xi = \alpha$. Puisque $s \leq \xi$ pour tout $s \in N_\xi$, on a $\alpha \leq \xi$ selon la définition de la borne supérieure. Supposons que $\alpha < \xi$. D'après 1.75, il y a un point rationnel $p \in (\alpha, \xi)$. Vu que $p < \xi$, on a $p \in N_\xi$, d'où $p \leq \sup N_\xi = \alpha$, ce qui contredit l'inclusion $p \in (\alpha, \xi)$. L'inégalité $\alpha < \xi$ est donc impossible, d'où $\alpha = \sup N_\xi = \xi$. D'une façon analogue on démontre que $\xi = \inf P_\xi$.

1.77. **S y s t è m e d é c i m a l d e n u m é r a t i o n d e s n o m b r e s r é e l s.** Nous vérifions ici que tout nombre réel peut être noté à l'aide de la suite des symboles 0, 1, 2, ..., 9.

Posons $9 + 1 = 10$.

Soit $\xi > 0$. D'après 1.72, il existe un (et seulement un) exposant p tel que

$$10^p \leq \xi < 10^{p+1}.$$

Choisissons ensuite un nombre θ_0 (de la suite 1, 2, ..., 9) tel que

$$\theta_0 \cdot 10^p \leq \xi < (\theta_0 + 1) \cdot 10^p.$$

Le nombre θ_0 est déterminé lui aussi d'une façon unique puisque les intervalles

$$0 \cdot 10^p \leq x < (0 + 1) \cdot 10^p,$$

pour différents $\theta = 0, 1, \dots, 9$, sont disjoints. Ensuite, étant donné θ_0 , choisissons un nombre θ_1 (de la suite 0, 1, 2, ..., 9) tel que

$$\theta_0 \cdot 10^p + \theta_1 \cdot 10^{p-1} \leq \xi < \theta_0 \cdot 10^p + (\theta_1 + 1) \cdot 10^{p-1}.$$

En continuant indéfiniment on obtient une suite de symboles (de chiffres de 0 à 9)

$$\theta_0 \theta_1 \theta_2 \dots \quad (\theta_0 \neq 0). \quad (1)$$

Pour fixer le nombre p on procède comme suit : si $p \geq 0$, on met une virgule entre les symboles θ_p et θ_{p+1} ; si $p < 0$, i.e. $p = -q$, $q > 0$, on place q zéros supplémentaires en tête de la suite (1) et on sépare le premier par une virgule.

Ainsi, nous avons fait correspondre à tout nombre réel ξ un symbole de la forme (1) précédé éventuellement de quelques zéros et avec une virgule à un certain endroit. Ce symbole est appelé *représentation décimale du nombre ξ* ; les chiffres $\theta_0, \theta_1, \dots$ dans leurs positions relatives dans la suite (1) sont appelés *signes décimaux du nombre ξ* . Pour le nombre 1 la représentation décimale a la forme 1,000...; elle est analogue pour les nombres 2, 3, ..., 9. Pour le nombre 10, la représentation décimale a la forme 10,000...

Pour les nombres de la forme $\frac{s}{10^t}$ (« fractions décimales ») avec s et t entiers non négatifs, et seulement pour eux, le nombre de décimales différentes de 0 est au plus égal à t .

Il est impossible que, dans (1), tous les chiffres à partir d'un certain soient les neuf. En effet, s'il n'y avait que des neuf à partir d'un numéro n après la virgule, ceci signifierait que le nombre ξ appartient aux intervalles

$$\begin{aligned} r + \frac{9}{10^n} &\leq \xi < r + \frac{10}{10^n} = r + \frac{1}{10^{n-1}}, \\ r + \frac{9}{10^n} + \frac{9}{10^{n+1}} &\leq \xi < r + \frac{9}{10^n} + \frac{10}{10^{n+1}} = r + \frac{1}{10^{n-1}}, \\ &\dots \dots \dots \\ r + \frac{9}{10^n} + \frac{9}{10^{n+1}} + \dots + \frac{9}{10^{n+h}} &\leq \xi < r + \frac{1}{10^{n-1}}, \\ &\dots \dots \dots \end{aligned}$$

or tous ces intervalles n'ont aucun point commun (1.74).

Réciproquement, soit une suite quelconque de chiffres de 0 à 9 :

$$\tau_1 \tau_2 \dots \quad (2)$$

avec une virgule à un certain endroit; de plus, supposons que les τ_i ne soient pas tous nuls et que les chiffres différents de 9 se rencontrent aussi loin que l'on veut. Montrons qu'il existe un nombre $\xi > 0$ pour lequel (2) se confond avec le symbole (1) représentant ce nombre.

Soit τ_m le premier chiffre différent de 0 dans (2). La virgule est placée soit $q \geq 0$ chiffres après τ_m (sans compter τ_m), soit $t \geq 1$ chiffres avant τ_m (y compris τ_m); dans le second cas, posons $q = -t$.

Maintenant posons

$$\xi = \sup_h \{ 10^q \cdot \tau_m + 10^{q-1} \cdot \tau_{m+1} + 10^{q-2} \cdot \tau_{m+2} + \dots + 10^{q-h} \cdot \tau_{m+h} \}$$

et montrons que le développement décimal de ce nombre ξ coïncide avec (2). Fixons un nombre naturel s , ensuite choisissons $r > s$ de façon à avoir $\tau_{m+r} \leq 8$, et soit un $k > r$. On trouve alors en som-

mant une progression géométrique que

$$\begin{aligned} 10^{q-(s+1)} \cdot \tau_{m+s+1} + \dots + 10^{q-r} \cdot \tau_{m+r} + \dots + 10^{q-h} \cdot \tau_{m+h} &\leq \\ &\leq 9 \cdot 10^{q-(s+1)} + \dots + 9 \cdot 10^{q-r} + \dots + 9 \cdot 10^{q-h} = 10^{q-r} = \\ &= 9 \cdot \frac{10^{q-(s+1)} - 10^{q-(h+1)}}{1 - 10^{-1}} = 10^{q-r} < 10^{q-s} - 10^{q-r}. \end{aligned}$$

C'est pourquoi

$$\begin{aligned} \xi &= \sup_k \{ 10^q \cdot \tau_m + \dots + 10^{q-s} \cdot \tau_{m+s} + \dots + 10^{q-h} \cdot \tau_{m+h} \} \leq \\ &\leq 10^q \cdot \tau_m + \dots + 10^{q-s} \cdot \tau_{m+s} + 10^{q-s} - 10^{q-r} < \\ &< 10^q \cdot \tau_m + \dots + 10^{q-s} (\tau_{m+s} + 1). \end{aligned}$$

Donc, pour tout $s = 0, 1, 2, \dots$, on a

$$\begin{aligned} 10^q \cdot \tau_m + \dots + 10^{q-s} \cdot \tau_{m+s} &\leq \xi < 10^q \cdot \tau_m + \dots \\ &\dots + 10^{q-s} (\tau_{m+s} + 1). \end{aligned}$$

En y posant $s = 0, 1, 2, \dots$ et en se rappelant la définition du nombre p et des signes décimaux $\theta_0, \theta_1, \dots$ du nombre ξ , on trouve $p = q$, $\theta_0 = \tau_m$, $\theta_1 = \tau_{m+1}$, $\dots$, d'où la coïncidence du développement décimal du nombre ξ avec le symbole (2).

Si $\xi < 0$, alors $-\xi > 0$, et par conséquent

$$-\xi = \tau_1 \tau_2 \dots$$

ce qu'on a montré plus haut; nous posons par définition

$$\xi = -\tau_1 \tau_2 \dots$$

Enfin, pour $\xi = 0$ nous posons

$$\xi = 0,0000 \dots$$

Ceci achève la construction du système décimal.

1.78. A la place du nombre 10, on aurait pu choisir un autre nombre entier $P > 1$. Le système correspondant de représentation des nombres réels est appelé système de base P . A côté du système décimal, on rencontre le plus souvent les systèmes *binaires* et *ternaires*, où P est respectivement 2 et 3. Le système binaire n'utilise que deux chiffres 0 et 1 pour représenter n'importe quel nombre réel, et le système ternaire n'a que les chiffres 0, 1, 2.

§ 1.8. Principe d'intervalles fermés emboîtés de Cantor

1.81. Soit un ensemble d'intervalles de l'axe réel R qui possède la propriété suivante: pour tout couple d'intervalles, l'un est inclus dans l'autre. Un tel ensemble d'intervalles sera appelé *système d'intervalles emboîtés*.

On a vu dans 1.74 que l'intersection d'un système d'intervalles emboîtés peut bien être vide. Les intervalles considérés dans 1.74 étaient semi-ouverts. A fortiori, un système d'intervalles ouverts emboîtés peut ne pas avoir de points communs. Néanmoins, si chacun des intervalles en question contient ses deux extrémités, c'est-à-dire est fermé, le point commun existe toujours; on a la proposition importante que voici:

Principe d'intervalles fermés emboîtés de Cantor. *Pour tout système Q d'intervalles fermés emboîtés $[a, b]$, il existe un point qui appartient à tous les intervalles du système. D'une façon plus stricte, il existe*

$$\xi = \sup \{a : [a, b] \in Q\};$$

$$\eta = \inf \{b : [a, b] \in Q\}.$$

Alors $\xi \leq \eta$, et l'intervalle $[\xi, \eta]$ est l'intersection de tous les intervalles du système Q .

Démonstration. Soient $E = \{a : [a, b] \in Q\}$ l'ensemble des extrémités gauches et $F = \{b : [a, b] \in Q\}$ l'ensemble des extrémités droites des intervalles du système Q . Pour n'importe quels deux intervalles $[a_1, b_1] \subset [a_2, b_2]$ du système Q , nous avons $a_2 \leq a_1 \leq b_1 \leq b_2$ de sorte que tout $a \in E$ ne dépasse aucun $b \in F$. Il résulte de 1.62.b que les nombres ξ et η figurant dans l'énoncé du théorème existent et vérifient l'inégalité $\xi \leq \eta$. Pour tout $[a, b] \in Q$ nous avons $a \leq \xi \leq \eta \leq b$, de sorte que $[a, b] \supset [\xi, \eta]$, d'où $\Pi [a, b] \supset [\xi, \eta]$. Il est aisé de voir que $\Pi [a, b]$ n'est composé que de points de l'intervalle $[\xi, \eta]$: pour tout point x n'appartenant pas à l'intervalle $[\xi, \eta]$, pour la raison par exemple que $x < \xi$, on peut trouver une extrémité gauche a pour laquelle $x < a < \xi = \sup \{a\}$, donc x n'appartient pas à l'intervalle correspondant $[a, b]$. Si $\xi = \eta$, alors en disant « intervalle $[\xi, \eta]$ » nous entendons le point $\xi = \eta$.

1.82. Dans quelles conditions l'intersection d'un système d'intervalles fermés emboîtés contient-elle un seul point? La réponse est donnée par le théorème suivant:

Théorème. *L'intersection d'un système Q d'intervalles fermés emboîtés est formée d'un seul point si, et seulement si, pour tout $\varepsilon > 0$, il y a dans le système Q un intervalle $[a, b]$ de longueur $b - a < \varepsilon$.*

Démonstration. D'après le principe de Cantor, les intervalles du système Q ont pour intersection un intervalle $[\xi, \eta]$

qui se réduit à un point si $\eta = \xi$. Et si $\eta \neq \xi$, la longueur de tout intervalle $[a, b] \supset \{\xi, \eta\}$ du système Q est au moins égale à $\eta - \xi$; par conséquent, s'il y a dans le système Q des intervalles de longueur arbitrairement petite, leur intersection est un seul point. Réciproquement, comme $\eta = \inf \{b : [a, b] \in Q\}$, $\xi = \sup \{a : [a, b] \in Q\}$, il existe dans le système Q , pour tout $\varepsilon > 0$ donné, un intervalle $[a_1, b_1]$ pour lequel $a_1 > \xi - \varepsilon/2$ et un intervalle $[a_2, b_2]$ pour lequel $b_2 < \eta + \varepsilon/2$. Si, par exemple, $[a_1, b_1] \supset [a_2, b_2]$, alors nous avons $a_2 \geq a_1 > \xi - \varepsilon/2$, $b_2 < \eta + \varepsilon/2$. Si $\xi = \eta$, on a $b_2 - a_2 < \varepsilon$, de sorte que, dans le système Q , il y a un intervalle de longueur $< \varepsilon$. Le théorème est démontré.

§ 1.9. Ensemble des nombres réels achevé

1.91. Définition. L'ensemble $\bar{R}$ des nombres réels achevé est composé de l'ensemble R de tous les nombres réels complété par deux symboles ou points $-\infty$ et ∞ (plus exactement $+\infty$).

La relation d'ordre s'étend sur ces symboles suivant la règle:

$$\begin{aligned} -\infty &< x && \text{pour tout } x \in R; \\ x &< \infty && \text{pour tout } x \in R; \\ -\infty &< \infty. \end{aligned}$$

Les axiomes d'ordre 1.23.a-d restent valables pour l'ensemble achevé. Les nombres réels ordinaires, contrairement aux symboles $-\infty$ et ∞ , sont dits *finis*.

1.92. Pour tout ensemble non vide $E \subset \bar{R}$, les quantités $\sup E$ et $\inf E$ sont définies par la règle suivante. Si E ne contient pas le point ∞ et est majoré (cf. 1.24), alors $\sup E$ conserve le même sens qu'en 1.24; dans les autres cas (i.e. si E contient ∞ ou s'il ne contient pas ∞ , mais n'est pas majoré) nous posons $\sup E = \infty$. D'une façon analogue, si E ne contient pas le point $-\infty$ et est minoré, le nombre $\inf E$ conserve le sens qu'on lui a attribué dans 1.61; dans les autres cas nous posons $\inf E = -\infty$. Ainsi, dans le système $\bar{R}$, tout ensemble non vide possède la borne supérieure et la borne inférieure.

1.93. Si $a \leq b$ sont deux points de $\bar{R}$, alors l'ensemble

$$[a, b] = \{x \in \bar{R} : a \leq x \leq b\}$$

ainsi que l'ensemble

$$(a, b) = \{x \in \bar{R} : a < x < b\}$$

s'appellent *intervalle* aux extrémités a et b .

1.94. Formulons l'extension du principe d'intervalles fermés emboîtés sur l'ensemble $\bar{R}$. Soit, sur la droite achevée $\bar{R}$, un ensemble Q d'intervalles emboîtés $[a, b]$ (au sens de 1.81). On affirme qu'ils contiennent tous un point $x \in \bar{R}$. En effet, soient A l'ensemble de toutes les extrémités gauches des intervalles du système Q et B l'ensemble des extrémités droites. Posons

$$\xi = \sup A, \quad \eta = \inf B.$$

De même que dans 1.81 on démontre que $\xi \leq \eta$ et que l'intersection de tous les intervalles $[a, b] \in Q$ est identique à l'intervalle $[\xi, \eta]$, donc non vide.

Exercices

1. Montrer que, quels que soient deux nombres réels x et y , on a l'inégalité

$$|x - y| \geq ||x| - |y||,$$

et pour n'importe quels $x, y_1, \dots, y_n$ l'inégalité

$$|x - y_1 - \dots - y_n| \geq ||x| - |y_1| - \dots - |y_n||.$$

2. On appelle *somme arithmétique* $A + B$ de deux ensembles numériques A et B l'ensemble de toutes les sommes $x + y$, où $x \in A, y \in B$. Démontrer que si A et B sont majorés, alors $A + B$ l'est aussi et

$$\sup(A + B) = \sup A + \sup B.$$

3. On appelle *produit arithmétique* AB de deux ensembles numériques A et B l'ensemble de tous les produits $x \cdot y$, où $x \in A, y \in B$. Démontrer que si A et B sont majorés et composés de nombres positifs, alors AB l'est aussi et

$$\sup(AB) = \sup A \cdot \sup B.$$

4. On appelle *puissance arithmétique* n -ième d'un ensemble A l'ensemble de tous les nombres x^n , où $x \in A$. (Notons qu'en général $A^n \neq A \cdot A$.) Démontrer que si A est composé de nombres positifs et majoré, alors A^n est aussi majoré et l'on a

$$\sup(A^n) = (\sup A)^n.$$

5. Si tout l'ensemble des nombres réels est partagé en deux ensembles disjoints non vides A et B tels que $a < b$ pour tous $a \in A, b \in B$, alors il existe un et seulement un nombre γ tel que $a \leq \gamma \leq b$ pour tous $a \in A$ et $b \in B$ (« coupure de Dedekind »).

Historique

Il paraît que la première théorie équivalente à la théorie moderne des nombres réels est construite par le mathématicien de la Grèce antique Eudoxe; elle est exposée dans les « Eléments » d'Euclide. En poursuivant pour l'essentiel

des buts purement logiques, cette théorie est peu appropriée aux calculs et constructions algébriques. Aussi les théories modernes remontent plutôt aux travaux de Gauss (1812), de Bolzano (1817) et surtout au « Cours d'analyse de l'école polytechnique » de Cauchy (1821) basé sur le principe d'intervalles fermés emboîtés que l'on tenait pour évident. Le pas suivant est réalisé vers 1872 par Dedekind, Weierstrass et Cantor qui, en suivant des voies différentes, définissent les nombres réels avec leurs propriétés en partant des nombres rationnels. Ainsi, la définition du système des nombres réels est ramenée à celle du système des nombres rationnels et, par la même, à celle du système des nombres naturels. C'est la raison pour laquelle Dedekind (1888), puis Peano (1891) formulent pour la première fois les axiomes des nombres naturels. Bien entendu, le problème se pose tout de suite de la non-contradiction du système d'axiomes de l'arithmétique que Hilbert (1900) inclut dans le nombre de problèmes mathématiques fondamentaux du XX^e siècle. Hilbert indique les moyens admissibles (« finis ») de sa résolution mais Gödel (1931) montre que le problème ne peut être résolu en principe avec ces moyens. La non-contradiction de l'arithmétique est démontrée par des moyens « transfinis » par Gentzen (1936) et Novikov (1943). La légitimité de ces moyens est loin d'être approuvée par tous les mathématiciens. Voir [4].

La première construction axiomatique directe de la théorie des nombres réels qui ne se base pas sur les nombres rationnels est proposée par Hilbert (1900).

Éléments de la théorie des ensembles

Le mérite immortel de Georg Cantor est de se hasarder dans le domaine de l'infini, sans craindre la lutte intérieure ni extérieure, non seulement avec les paradoxes imaginaires, les préjugés largement répandus, les sentences des philosophes, mais aussi avec le parti pris du nombre de grands mathématiciens. Ainsi il créa une nouvelle science, la théorie des ensembles.

F. Hausdorff (1927)

Les mathématiques (du gr. *mathema*, connaissance). Science qui a pour objet les structures mathématiques (ensembles dont les éléments sont en certaines relations).

Dictionnaire philosophique (ELP, 1968)

§ 2.1. Opérations sur les ensembles

2.11. Nous considérons ici, un peu plus en détail, les opérations sur les ensembles introduites dans 1.12. Rappelons les définitions de ces opérations. S'il y a des ensembles $A, B, C, \dots$, alors l'ensemble de tous les éléments appartenant à l'un au moins de ces ensembles s'appelle *réunion* des ensembles $A, B, C, \dots$. L'ensemble de tous les éléments appartenant à *chacun* des ensembles $A, B, C, \dots$ à la fois s'appelle *intersection* des ensembles $A, B, C, \dots$.

La réunion S des ensembles $A, B, C, \dots$ est parfois appelée *somme* et notée $S = A + B + C + \dots$; l'intersection D est appelée *produit* et notée $D = ABC \dots$ *). Il y a certaines raisons d'employer ces termes « arithmétiques ». Par exemple, pour trois ensembles quelconques A, B, C , on a l'égalité

$$(A + B)C = AC + BC.$$

Rappelons que deux ensembles sont égaux si tout élément de l'un est élément de l'autre. Donnons la démonstration

*) Ne pas confondre avec la « somme arithmétique » et le « produit arithmétique » introduits dans les exercices pour le chapitre 1.

de l'égalité ci-dessus comme exemple simple mais typique de raisonnements sur l'égalité des ensembles.

Nous avons à montrer que tout élément x appartenant à $(A + B) C$ (premier membre) appartient à $AC + BC$ (second membre) et, réciproquement, tout élément y appartenant à $AC + BC$ appartient aussi à $(A + B) C$. Supposons d'abord qu'un x appartienne à $(A + B) C$. En tant qu'élément de l'intersection des ensembles $A + B$ et C , l'élément x doit appartenir à chacun d'eux : par conséquent, nous avons :

$$x \in A + B \quad \text{et} \quad x \in C.$$

Puisque x appartient à la réunion de A et B , il appartient à l'un au moins des termes, par exemple à A . Or les inclusions $x \in A$, $x \in C$ impliquent $x \in AC$, d'où $x \in AC + BC$. Si x appartient non pas à A mais à B , on obtient de même $x \in BC$. $x \in AC + BC$, ce qu'il nous fallait. Réciproquement, si y appartient à la somme $AC + BC$, il appartient à l'un des termes, par exemple $y \in BC$. Mais alors $y \in B$ et $y \in C$; ensuite, $y \in B$ implique $y \in A + B$ et, définitivement, $y \in (A + B) C$. Le cas $y \in AC$ est traité d'une façon analogue, ce qui achève la démonstration.

Il faut cependant signaler qu'on est loin de pouvoir étendre toutes les règles arithmétiques aux opérations sur les ensembles. Par exemple, pour trois ensembles A , B , C , on a les formules :

$$A + A = A,$$

$$A \cdot A = A,$$

$$A + BC = (A + B) (A + C),$$

qui ne ressemblent plus à des égalités arithmétiques habituelles. Nous proposons au lecteur de les prouver.

2.12. Introduisons maintenant une nouvelle opération.

Si B est un sous-ensemble d'un ensemble A , alors l'ensemble de tous les éléments de A qui n'appartiennent pas à B s'appelle *complémentaire de l'ensemble B par rapport à l'ensemble A* et est désigné par CB ou $A - B$.

Signalons les formules évidentes :

$$C(CB) \equiv A - (A - B) = B,$$

$$(A - B) + B = A.$$

Notons que, pour deux ensembles arbitraires A et B , la formule

$$(A + B) - B = A$$

est, en général, fausse; elle n'est juste que dans le cas où A et B sont disjoints.

Parmi les formules plus compliquées, signalons l'une que nous rencontrerons souvent dans la suite :

$$C \sum_v B_v = \prod_v C B_v. \quad (1)$$

Elle se lit : *le complémentaire de la réunion de quelques ensembles est l'intersection de leurs complémentaires.*

Démontrons cette formule. Soit $x \in C \sum_v B_v$; alors $x \notin \sum_v B_v$; cela veut dire que, pour tout v , on a $x \notin B_v$, i.e. $x \in C B_v$; d'où $x \in \prod_v C B_v$. Réciproquement, si $x \in \prod_v C B_v$, alors $x \in C B_v$ pour tout v , c'est-à-dire que $x \notin B_v$ pour tout v ; alors $x \notin \sum_v B_v$, donc $x \in C \sum_v B_v$, ce qu'il fallait démontrer.

En prenant encore une fois les complémentaires de deux membres de l'égalité (1) et en posant $A_v = C B_v$, on obtient la formule

$$\sum_v C A_v = C \prod_v A_v, \quad (2)$$

c'est-à-dire que *le complémentaire de l'intersection de quelques ensembles est la réunion de leurs complémentaires.*

On peut résumer les résultats obtenus dans une règle générale : *en permutant le symbole C avec $\sum$ (ou $\prod$) on doit remplacer $\sum$ par $\prod$ (respectivement $\prod$ par $\sum$).*

§ 2.2. Equipotence d'ensembles

Nous voulons maintenant établir des règles permettant de comparer les ensembles du point de vue de la quantité d'éléments.

Pour des ensembles finis, le problème ne se pose pas : en comptant les éléments de chacun de deux ensembles finis A et B on peut voir directement lequel d'entre eux est le plus riche en éléments. Il est naturel de dire que deux ensembles finis A et B sont équipotents s'ils ont un même nombre d'éléments. Néanmoins, cette définition de l'équipotence ne s'étend pas directement au cas d'ensembles infinis. Mettons-la sous la forme qui permet une telle extension. Notons que pour l'établissement de l'équipotence ou de la non-équipotence des ensembles finis A et B il n'est pas nécessaire de compter réellement leurs éléments. Si, par exemple, l'ensemble A est l'ensemble d'élèves d'une salle d'études et B est l'ensemble des chaises de la même salle, alors, au lieu de compter les élèves et les chaises séparément, il vaut mieux inviter chaque élève à occuper l'une des chaises disponibles et il sera clair sans calcul si les ensembles en question sont équipotents ou non.

Sous sa forme abstraite, le procédé qu'on a utilisé dans cet exemple, est l'établissement d'une correspondance biunivoque entre les ensembles A et B .

2.21. Définition. Si à tout élément d'un ensemble A on a fait correspondre, d'une certaine façon, un élément d'un ensemble B , de sorte que tout élément de l'ensemble B corresponde à un, et seulement à un, élément de l'ensemble A , alors on dit qu'il y a une *correspondance biunivoque* entre les ensembles A et B , et ceux-ci sont dits *équipotents*.

Cette nouvelle définition de l'équipotence convient pour n'importe quels ensembles, pas obligatoirement finis; ainsi, l'ensemble infini A des nombres naturels $1, 2, \dots$ est équipotent à l'ensemble B des nombres entiers négatifs $-1, -2, \dots$, et la correspondance biunivoque entre les ensembles A et B s'établit d'après la règle: à tout nombre $n \in A$ on fait correspondre le nombre $-n \in B$.

De même, l'ensemble des nombres naturels $1, 2, \dots$ est équipotent à l'ensemble de tous les nombres positifs pairs $2, 4, \dots$; la correspondance se réalise suivant la règle $n \rightarrow 2n$.

On voit sur cet exemple qu'un ensemble peut être équipotent à l'un de ses sous-ensembles propres; cette situation ne peut, bien entendu, se réaliser que pour des ensembles infinis. Une correspondance entre les éléments $x \in A$ et $y \in B$ est désignée par $x \leftrightarrow y$.

La relation d'équipotence est désignée par le signe $\sim$. Il est aisé de voir que cette relation est *réflexive* (i.e. $A \sim A$), *symétrique* (si $A \sim B$, alors $B \sim A$) et *transitive* (si $A \sim B$ et $B \sim C$, alors $A \sim C$). Si deux ensembles sont équipotents, on dit aussi qu'ils ont même puissance.

2.22. Exemples. L'intervalle $[0, a]$, $a > 0$, est équipotent à l'intervalle $[0, 1]$; la correspondance s'établit par la formule $x \in [0, 1] \leftrightarrow y = ax \in [0, a]$. Deux intervalles quelconques $[a, b]$ et $[a + h, b + h]$ de même longueur $b - a$ sont équipotents; la correspondance s'établit par la formule $x \in [a, b] \leftrightarrow y = x + h \in [a + h, b + h]$. Donc, en général, n'importe quels deux intervalles fermés de la droite numérique sont équipotents. De même sont équipotents n'importe quels deux intervalles ouverts de la droite numérique. La correspondance $x \leftrightarrow y = \frac{1}{x}$ établit l'équipotence de l'intervalle $0 < x < 1$ et du demi-axe $y > 1$, la correspondance $x \leftrightarrow y = \frac{x}{1+x^2}$ celle de toute la droite numérique et de l'intervalle $(-1, 1)$.

Démontrons à présent que l'intervalle $[a, b]$ est équipotent à l'intervalle (a, b) . Considérons une suite quelconque A de points de l'intervalle $[a, b]$ contenant ses deux extrémités: $x_1 = a$, $x_2 = b$, $x_3, \dots, x_n, \dots$. Alors $x_3, x_4, \dots$ appartiennent à l'inter-

valle (a, b) de même que les autres points de l'intervalle $[a, b]$ non inclus dans la suite considérée. Établissons la correspondance entre les points de l'intervalle fermé et ceux de l'intervalle ouvert d'après la règle suivante :

$$\begin{aligned} x_1 &\leftrightarrow x_3, \\ x_2 &\leftrightarrow x_4, \\ &\vdots \\ x_n &\leftrightarrow x_{n+2}, \\ &\vdots \\ y \notin A &\leftrightarrow y. \end{aligned}$$

Cette correspondance est bien biunivoque. Donc, les intervalles $[a, b]$ et (a, b) sont équipotents. En appliquant les correspondances obtenues plus haut, nous concluons que *tout intervalle ouvert est équipotent à tout intervalle fermé, à la demi-droite et à la droite numérique.*

§ 2.3. Ensembles dénombrables

2.31. Définition. Un ensemble équipotent à l'ensemble de tous les nombres naturels $1, 2, \dots$ s'appelle *ensemble dénombrable*.

Autrement : un ensemble est dénombrable si tous ses éléments peuvent être numérotés avec tous les nombres naturels. Donnons quelques théorèmes simples sur les ensembles dénombrables.

2.32. Tout sous-ensemble infini B d'un ensemble dénombrable A est lui-même dénombrable. En effet, on peut numérotter les éléments de l'ensemble B dans l'ordre de leur succession dans A (et l'on sera forcé d'utiliser tous les entiers naturels car B est infini).

2.33. La réunion d'une famille finie ou dénombrable d'ensembles dénombrables est un ensemble dénombrable.

Démonstration. Considérons d'abord le cas de deux ensembles. Soient $A = (a_1, a_2, \dots)$ et $B = (b_1, b_2, \dots)$ deux ensembles dénombrables. Ordonnons tous les éléments de ces deux ensembles d'après la règle suivante :

$$a_1, b_1, a_2, b_2, a_3, b_3, \dots$$

Maintenant on peut numérotter tous ces éléments par l'ordre de leur succession dans cette suite. Tout élément rencontré deux fois (qui appartient donc à A et B) sera muni d'un indice la première fois et omis la deuxième. Définitivement, tout élément de la réunion des ensembles A et B aura son numéro, ce que nous cherchions.

Ainsi, l'ensemble de tous les nombres entiers $0, \pm 1, \pm 2, \dots$ est dénombrable comme réunion de deux ensembles dénombrables $1, 2, 3, \dots$ et $0, -1, -2, \dots$.

D'une façon analogue, on démontre le théorème dans le cas de trois, quatre et, en général, d'un nombre fini d'ensembles dénombrables. Pour le cas d'une famille dénombrable d'ensembles dénombrables

$$\begin{aligned} A_1 &= \{a_{11}, a_{12}, \dots, a_{1n}, \dots\}, \\ A_2 &= \{a_{21}, a_{22}, \dots, a_{2n}, \dots\}, \\ &\vdots \\ A_k &= \{a_{k1}, a_{k2}, \dots, a_{kn}, \dots\}, \\ &\vdots \end{aligned}$$

la seule différence est dans la manière de ranger les éléments de tous ces ensembles dans une suite; on les dispose, par exemple, par groupes d'éléments ayant une même somme d'indices:

$$a_{11}; a_{21}, a_{12}; a_{31}, a_{22}, a_{13}; a_{41}, a_{32}, a_{23}, a_{14}; \dots,$$

la suite de la démonstration ne varie pas.

2.34. *L'ensemble de tous les nombres rationnels (c'est-à-dire les nombres de la forme $\frac{p}{q}$, p et q étant entiers) est dénombrable.*

En effet, l'ensemble de tous les nombres rationnels est la réunion des ensembles dénombrables suivants:

- 1) ensemble A_1 de tous les nombres entiers $n = 0, \pm 1, \pm 2, \dots$;
- 2) ensemble A_2 des fractions de la forme $\frac{n}{2}$, $n = 0, \pm 1, \pm 2, \dots$;
- 3) ensemble A_3 des fractions de la forme $\frac{n}{3}$, $n = 0, \pm 1, \pm 2, \dots$;
- $\vdots$
- k) ensemble A_k des fractions de la forme $\frac{n}{k}$, $n = 0, \pm 1, \pm 2, \dots$;
- $\vdots$

Les $A_1, A_2, \dots, A_k, \dots$ forment une famille dénombrable d'ensembles; chacun d'eux étant dénombrable, leur réunion l'est aussi en vertu de 2.33, ce qu'il fallait démontrer.

2.35. *Si*

$$A = (a_1, a_2, \dots, a_k, \dots) \text{ et } B = (b_1, b_2, \dots, b_n, \dots)$$

sont des ensembles dénombrables, alors l'ensemble de tous les couples (a_k, b_n) ($k, n = 1, 2, \dots$) est aussi dénombrable.

En effet, l'ensemble de tous ces couples peut être représenté sous forme de réunion de la famille dénombrable des ensembles

dénombrables

$$A_1 = \{(a_1, b_1), (a_1, b_2), \dots, (a_1, b_n), \dots\},$$

$$A_2 = \{(a_2, b_1), (a_2, b_2), \dots, (a_2, b_n), \dots\},$$

• • • • •

$$A_k = \{(a_k, b_1), (a_k, b_2), \dots, (a_k, b_n), \dots\},$$

• • • • •

et, en vertu de 2.33, est aussi un ensemble dénombrable.

Cet exemple admet une interprétation géométrique: à tout couple (a_n, b_n) il correspond un point du plan de coordonnées a_n, b_n ; on voit, en particulier, que l'ensemble de tous les points du plan dont les coordonnées sont rationnelles est dénombrable.

2.36. L'ensemble de tous les polynômes $P(x) = a_0 + a_1x + \dots + a_nx^n$ (de degrés quelconques) à coefficients rationnels $a_0, a_1, \dots, a_n$ est dénombrable.

L'ensemble des polynômes de ladite forme est la réunion de la famille dénombrable des ensembles A_n ($n = 0, 1, 2, \dots$), où A_n désigne l'ensemble des polynômes de degrés $\leq n$. Par conséquent, vu le théorème 2.33, il suffit de montrer que chacun des ensembles A_n est dénombrable. Pour $n = 0$, il s'agit de la dénombrabilité de l'ensemble des nombres rationnels qui a été établie dans 2.34. Ensuite, on procède par récurrence: supposons que la dénombrabilité d'un ensemble A_n soit établie et démontrons celle de l'ensemble A_{n+1} .

Tout élément de l'ensemble A_{n+1} peut être mis sous la forme

$$O(x) \leq a_{n+1}x^{n+1}.$$

où $Q(x)$ est un polynôme d'un degré $\leq n$ à coefficients rationnels, donc un élément de l'ensemble A_n , et a_{n+1} un nombre rationnel.

L'ensemble des polynômes $Q(x)$ est dénombrable par hypothèse, l'ensemble des nombres a_{n+1} l'est également. Donc, à tout élément de l'ensemble A_{n+1} , on peut faire correspondre un couple $(Q(x), a_{n+1})$ dont chaque composante parcourt un ensemble dénombrable de valeurs. Vu 2.35, l'ensemble A_{n+1} est aussi dénombrable, ce qu'il fallait démontrer.

2.37. L'ensemble de tous les nombres algébriques (c'est-à-dire des racines des polynômes à coefficients rationnels) est dénombrable.

D'après 2.36, tous les polynômes à coefficients rationnels peuvent être numérotés à l'aide des nombres naturels, de sorte que ces polynômes forment une suite

$$P_1(x), P_2(x), \dots, P_n(x), \dots$$

Or, chacun des polynômes mentionnés possède un nombre fini de racines (cf. 5.87). En écrivant d'abord toutes les racines du polynôme $P_1(x)$, ensuite celles du polynôme $P_2(x)$, etc., nous arrivons à la possibilité de numérotter tous les nombres algébriques.

§ 2.4. Ensembles ayant la puissance du continu

Il existe des ensembles infinis dont les éléments ne peuvent être numérotés avec les nombres naturels. De tels ensembles sont dits *non dénombrables*. Un *continu*, c'est-à-dire l'ensemble de tous les points d'un intervalle, en sert d'exemple.

2.41. Théorème (Cantor, 1874). *L'ensemble de tous les points de l'intervalle $0 \leq x \leq 1$ est non dénombrable.*

Démonstration. Supposons que, au contraire, l'ensemble de tous les points de l'intervalle $[0, 1]$ soit dénombrable, de sorte qu'ils forment une suite $x_1, x_2, \dots, x_n, \dots$. Cette suite étant, construisons une suite d'intervalles emboîtés.

Partageons l'intervalle $[0, 1]$ en trois parties égales. Où que soit un point x_1 , il ne peut appartenir à la fois aux trois intervalles $\left[0, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, 1\right]$, et l'on peut donc en choisir un qui ne contient pas le point x_1 (ni à l'intérieur, ni sur la frontière); on désigne cet intervalle par Δ_1 . Ensuite, désignons par Δ_2 celle des trois parties égales de l'intervalle Δ_1 qui ne contient pas le point x_2 . Une fois les intervalles $\Delta_1 \supset \Delta_2 \supset \dots \supset \Delta_n$ ainsi construits, on désigne par Δ_{n+1} celle des trois parties égales de l'intervalle Δ_n qui ne contient pas le point x_{n+1} , etc. La suite infinie des intervalles $\Delta_1 \supset \Delta_2 \supset \dots$, en vertu du théorème 1.81, a un point commun ξ . Ce point ξ appartient à chacun des intervalles Δ_n et ne peut donc coïncider avec aucun point x_n . Cela montre que la suite $x_1, x_2, \dots, x_n, \dots$ n'épuise pas tous les points de l'intervalle $[0, 1]$, en contradiction avec l'hypothèse. Le théorème est démontré.

2.42. Nous avons vu que tous les nombres rationnels de l'intervalle $[0, 1]$ forment un ensemble dénombrable. Les autres nombres de l'intervalle sont dits *irrationnels*. Nous voyons maintenant que les nombres irrationnels sont « beaucoup plus nombreux » que les nombres rationnels: d'une façon plus rigoureuse, les nombres irrationnels forment un ensemble nécessairement non dénombrable (sinon, l'ensemble de tous les nombres de l'intervalle $0 \leq x \leq 1$ serait dénombrable comme réunion de deux ensembles dénombrables). Qui plus est, puisque les nombres irrationnels algébriques (racines de polynômes à coefficients rationnels) forment un ensemble dénombrable (2.37), l'ensemble des nombres *transcendants*, c'est-

à-dire des nombres qui ne sont pas des racines de polynômes à coefficients rationnels, est non dénombrable.

A propos, le raisonnement précédent démontre l'existence même de nombres transcendants, qui n'est nullement évidente a priori.

2.43. Définition. Tout ensemble équipotent à l'ensemble de tous les points de l'intervalle $[0, 1]$ est appelé *ensemble ayant la puissance du continu*.

On a vu que les ensembles des points de tout intervalle fermé $[a, b]$, de tout intervalle ouvert (α, β) et enfin de toute la droite $-\infty < x < \infty$ sont tous équipotents à l'ensemble des points de l'intervalle $[0, 1]$; donc ils ont tous la puissance du continu.

§ 2.5. Notion de structure mathématique. Isomorphisme de structures

2.51. D'un point de vue général, les mathématiques n'ont affaire qu'à des ensembles. Mais la richesse de telle ou telle théorie mathématique dépend de liens entre les éléments (et les sous-ensembles) des ensembles étudiés par la théorie. Ces liens sont formulés d'une façon abstraite, à l'aide d'axiomes. A titre d'exemple on peut citer le système des nombres réels considéré dans le chapitre 1: il représente un ensemble dont les éléments satisfont à un certain système d'axiomes, assez compliqué. Un ensemble muni de certaines conditions imposées à ses éléments et ses sous-ensembles est appelé *structure mathématique*. Une définition rigoureuse de la structure mathématique devrait bien comprendre une définition générale desdites conditions, mais nous préférons de ne pas insister là-dessus [15] en nous bornant à une description et à quelques exemples.

2.52. Deux structures avec des conditions d'un même type sont dites *isomorphes* si l'on peut établir une correspondance biunivoque entre leurs éléments de façon que la satisfaction ou la non-satisfaction des conditions déterminant la structure soit conservée. Toute structure est isomorphe à elle-même: l'application identique est, évidemment, bijective *) et conserve toutes les conditions imposées à des éléments et des sous-ensembles de la structure. Mais il existe aussi des applications non identiques d'une structure sur elle-même qui conservent les conditions décrivant la structure; de telles applications s'appellent *automorphismes* de la structure.

Considérons, à titre d'exemple, la structure d'ensemble *totallement ordonné*. On appelle ainsi un ensemble E des éléments $x, y, \dots$

*) L'existence de l'application bijective de l'ensemble A sur l'ensemble B entraîne celle de l'application réciproque (bijective elle aussi) de B sur A , ce qui est équivalent à l'existence de la correspondance biunivoque entre les éléments de ces deux ensembles. (N.D.R.)

muni de la condition supplémentaire que, pour deux éléments différents quelconques x, y , il est établi l'une des relations $x < y$ (« plus petit que » ou « inférieur à ») ou $y < x$; de plus, si $x < y$ et $y < z$, alors $x < z$. La droite numérique (achevée) ou tout sous-ensemble de la droite numérique avec la relation d'ordre habituelle en servent d'exemples. Deux ensembles ordonnés E_1, E_2 sont dits isomorphes si une correspondance biunivoque est établie entre eux, de sorte que $x_1 \in E_1, y_1 \in E_1, x_1 < y_1$ impliquent $x_2 < y_2$ pour les éléments respectifs $x_2 \in E_2$ et $y_2 \in E_2$. En ce sens, toute la droite numérique non achevée est isomorphe à n'importe quel intervalle ouvert; mais elle n'est pas isomorphe à un intervalle fermé (un intervalle fermé possède un élément minimal et l'axe non). Un automorphisme d'un ensemble totalement ordonné est n'importe quelle application bijective de cet ensemble sur lui-même qui conserve l'ordre: si $x_1 < y_1$, x_1 correspond à x_2 , y_1 correspond à y_2 , alors $x_2 < y_2$. Pour l'intervalle $(0, 1)$ en tant qu'ensemble totalement ordonné, la formule $y = x^n$, pour tout n fixe, détermine un automorphisme.

Il existe des structures dont le système d'axiomes définit la structure à un isomorphisme près; on d'autres termes, il est possible que deux structures quelconques munies d'un système d'axiomes donné soient isomorphes. On dit dans ce cas que le système d'axiomes d'une structure donnée est *complet* et que la structure est *unique*.

2.53. En tant qu'exemple, considérons le problème d'unicité du système des nombres réels. Dans le chapitre 1, nous avons appelé système des nombres réels un ensemble d'objets qui satisfont aux axiomes 1.21-1.24.

Une question se pose: est-ce que le système d'objets vérifiant tous ces axiomes est unique? Deux structures des nombres réels sont-elles isomorphes? D'une façon plus stricte: s'il y a deux exemplaires du système des nombres réels, peut-on établir une correspondance biunivoque entre leurs éléments de manière qu'aux résultats de l'addition et de la multiplication de nombres du premier système il corresponde les résultats des mêmes opérations sur les nombres correspondants du second système, que la relation d'ordre entre les éléments correspondants soit conservée et que les bornes des ensembles correspondants se correspondent? Le théorème suivant dont la démonstration sera faite dans 2.54-2.58 répond par l'affirmative.

Théorème. Soient R_1 et R_2 deux exemplaires du système des nombres réels; les nombres du premier système seront munis d'indice 1, ceux du deuxième d'indice 2. Il est alors possible d'établir une correspondance biunivoque entre les nombres $x_1 \in R_1$ et $x_2 \in R_2$ désignée par le symbole $\sim$ et telle que les propriétés suivantes soient valables:

1) $x_1 \sim x_2, y_1 \sim y_2$ entraînent $x_1 + y_1 \sim x_2 + y_2$;

- 2) $x_1 \sim x_2, y_1 \sim y_2$ entraînent $x_1 y_1 \sim x_2 y_2$;
 3) $x_1 \sim x_2, y_1 \sim y_2, x_1 < y_1$ entraînent $x_2 < y_2$;
 4) si un ensemble $A_1 \subset R_1$ est majoré, alors l'ensemble A_2 des éléments correspondants de R_2 l'est aussi et, qui plus est,

$$\sup A_1 \sim \sup A_2.$$

2.54. En abordant la démonstration du théorème 2.53 nous notons que l'équipotence des systèmes R_1 et R_2 au sens de la théorie des ensembles découle déjà de la possibilité de mettre les nombres de chaque système sous la forme de développements décimaux infinis (1.77). Mais nous employerons un procédé plus propre à vérifier les propriétés 1)-4). Désignons par Q_1 et Q_2 les ensembles des nombres rationnels des systèmes R_1 et R_2 respectivement. Une correspondance biunivoque naturelle s'établit entre Q_1 et Q_2 d'après la représentation même des nombres rationnels sous la forme $\frac{m}{n}$. Evidemment, les propriétés 1)-3) sont alors remplies puisque les opérations mentionnées s'expriment directement par les symboles m et n . Ensuite, si $\xi_1 \in R_1$ est un nombre quelconque, on a selon 1.76 :

$$\xi_1 = \sup N_1(\xi_1) = \inf P_1(\xi_1),$$

où $N_1(\xi_1)$ désigne l'ensemble des nombres rationnels $r_1 \leq \xi_1$ et $P_1(\xi_1)$ celui des nombres rationnels $s_1 \geq \xi_1$; ici toujours $r_1 \leq s_1$. Désignons par N_2 l'ensemble des nombres rationnels du système R_2 qui correspondent aux nombres rationnels de l'ensemble $N_1(\xi_1)$ (c'est-à-dire qui ont les mêmes notations $\frac{m}{n}$), et par P_2 l'ensemble des nombres rationnels du système R_2 qui correspondent aux nombres de l'ensemble $P_1(\xi_1)$. Nous avons $r_2 \leq s_2$ quels que soient $r_2 \in N_2$ et $s_2 \in P_2$, d'où $\sup N_2 \leq \inf P_2$ d'après 1.62.b. Tous les nombres rationnels du système R_2 appartenant soit à N_2 , soit à P_2 , il n'y a pas de nombres rationnels entre les nombres $\alpha_2 = \sup N_2$ et $\beta_2 = \inf P_2$. Or ceci veut dire que $\alpha_2 = \beta_2$, sinon il existerait un point rationnel entre α_2 et β_2 d'après 1.75. Faisons maintenant correspondre au nombre $\xi_1 \in R_1$ le nombre $\xi_2 = \alpha_2 = \beta_2 \in R_2$. Cette correspondance est biunivoque par symétrie.

2.55. Examinons maintenant les propriétés 3) et 4). Si $x_1 < y_1$, on choisit, à l'aide de 1.75, un nombre rationnel $r_1, x_1 < r_1 < y_1$; pour les nombres correspondants x_2, r_2, y_2 , selon la définition même de la correspondance, on a $x_2 < r_2 < y_2$, de sorte que 3) est vérifiée. Soit un A_1 majoré, $\xi_1 = \sup A_1$, et soient A_2 l'ensemble correspondant à A_1 et ξ_2 le nombre correspondant à ξ_1 . Nous avons $\xi_1 \geq x_1$ pour tout $x_1 \in A_1$, d'où $\xi_2 \geq x_2$ pour tout $x_2 \in A_2$, donc $\xi_2 \geq \sup A_2$. Si $\sup A_2 < \xi_2$, alors il y a un nombre y_2 (même rationnel) tel que $x_2 < y_2 < \xi_2$ pour tout $x_2 \in A_2$, d'où $x_1 < y_1 < \xi_1$ pour tout $x_1 \in A_1$ et pour y_1 qui correspond à y_2 . Mais alors $\sup A_1 \leq y_1 < \xi_1$,

en contradiction avec l'hypothèse. Ainsi $\sup A_2 = \xi_2$, ce qui prouve 4).

2.56. Pour démontrer les propriétés 1) et 2), établissons deux lemmes.

a. *L e m m e.* Soient des nombres réels y et z et un nombre rationnel $r > y + z$; il existe des nombres rationnels $u > y$, $v > z$ tels que $u + v = r$.

D é m o n s t r a t i o n. Soit $r - (y + z) = h$. Choisissons pour u n'importe quel nombre rationnel de l'intervalle $(y, y + h)$ et posons $v = r - u$; évidemment $v > r - (y + h) = z$, ce qu'il fallait démontrer.

b. *L e m m e.* Soient des nombres positifs y et z et un nombre rationnel $r > yz$; il existe des nombres rationnels $u > y$, $v > z$ tels que $uv = r$.

D é m o n s t r a t i o n. Soit $\frac{r}{yz} = h$. Choisissons pour u n'importe quel nombre rationnel de l'intervalle (y, yh) et posons $v = \frac{r}{u}$; évidemment, $v > \frac{r}{yh} = z$, ce qu'il nous fallait.

2.57. Vérifions à présent la propriété 1). Soient $x_1 \sim x_2$, $y_1 \sim y_2$, $x_1 + y_1 \sim x_2 + y_2$ et soit, par exemple, $z_2 > x_2 + y_2$. Choisissons un nombre rationnel $t_2 \in (x_2 + y_2, z_2)$. D'après le lemme 2.56.a il existe des nombres rationnels $u_2 > x_2$, $v_2 > y_2$, $u_2 + v_2 = t_2$. Pour les nombres correspondants z_1 , t_1 , u_1 , v_1 , nous avons $x_1 < u_1$, $y_1 < v_1$, $x_1 + y_1 < u_1 + v_1 = t_1$. Or $t_2 < z_2 \sim x_1 + y_1$, donc $t_1 < x_1 + y_1$ en contradiction avec ce qui précède. L'hypothèse $z_2 < x_2 + y_2$ tombe en défaut d'une façon analogue.

2.58. En passant à la propriété 2) envisageons-la d'abord pour x_1 et y_1 positifs; il leur correspond, selon 3), x_2 et y_2 également positifs. Soient $x_1 y_1 \sim x_2 y_2$ et, par exemple, $z_2 > x_2 y_2$. Choisissons un nombre rationnel $t_2 \in (x_2 y_2, z_2)$. D'après le lemme 2.56.b, il existe des nombres rationnels $u_2 > x_2$, $v_2 > y_2$, $u_2 v_2 = t_2$. Pour les nombres correspondants z_1 , t_1 , u_1 , v_1 , nous avons $x_1 < u_1$, $y_1 < v_1$, $x_1 y_1 < u_1 v_1 = t_1$. Or $t_2 < z_2 \sim x_1 y_1$, donc $t_1 < x_1 y_1$ en contradiction avec ce qui précède. L'hypothèse $z_2 < x_2 y_2$ tombe en défaut d'une façon analogue. Le cas où l'un des facteurs est nul est immédiat.

En passant au cas général et vu la règle des signes, il suffit de prouver que $-1 \cdot x_1 \sim -1 \cdot x_2$. Supposons que $-1 \cdot x_1$ corresponde à un z_2 . D'après 1), l'égalité $-1 \cdot x_1 + x_1 = 0$ implique $z_2 + x_2 = 0$, d'où $z_2 = -1 \cdot x_2$. Ainsi, notre théorème est complètement démontré.

2.59. Nous venons de montrer l'unicité, à un isomorphisme près, du système des nombres réels. Étudions la nature des automorphismes

de ce système. D'après 2.52, nous appelons automorphisme du système R une application bijective $R \rightarrow R$ qui conserve les résultats de l'addition et de la multiplication, la relation d'ordre et la borne supérieure. Montrons qu'une telle application $R \rightarrow R$ ne peut être qu'identique. Désignons par x' l'élément de R en lequel est transformé par l'automorphisme considéré un élément x . L'élément $1'$ est l'unité du corps R puisque, pour tout x , nous avons $1' \cdot x' = x'$; en vertu de l'unicité de l'unité (1.41.a), nous avons $1' = 1$. Il en résulte que $2' = 1' + 1' = 2$, ..., $n' = (n-1)' + 1' = n$;

$\left(\frac{m}{n}\right)' = \frac{m'}{n'} = \frac{m}{n}$. L'unicité de l'élément opposé (1.32) implique

$\left(-\frac{m}{n}\right)' = -\frac{m'}{n'} = -\frac{m}{n}$. Ainsi, l'automorphisme en question conserve le

système des nombres rationnels. Mais alors il conserve tout nombre réel car, en vertu de 1.76, nous avons pour tout $\xi \in R$:

$$\xi = \sup N_\xi,$$

où N_ξ est la classe des nombres rationnels qui ne dépassent pas ξ , et l'automorphisme considéré conserve la relation d'ordre et la borne supérieure. Donc, cet automorphisme est l'application identique $\xi \rightarrow \xi$.

§ 2.6. Espace n -dimensionnel

2.61. Définition. Soit un nombre naturel n ; désignons par R_n l'ensemble de tous les complexes (« vecteurs ») de n nombres réels

$$x = (x_1, \dots, x_n).$$

Les nombres $x_1, \dots, x_n$ s'appellent *coordonnées* du vecteur x . Nous dirons qu'un vecteur $x = (x_1, \dots, x_n)$ est égal à un vecteur $y = (y_1, \dots, y_n)$ si, et seulement si, les égalités $x_1 = y_1, \dots, x_n = y_n$ ont lieu.

L'ensemble R_n est appelé *espace réel n -dimensionnel*.

2.62. Définissons l'opération d'addition dans l'espace R_n par la formule

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n). \quad (1)$$

Montrons que les axiomes d'addition 1.21 sont remplis où les symboles $x, y, \dots$ désignent déjà les éléments de l'espace R_n . Soient $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n), z = (z_1, \dots, z_n)$. Nous avons:

$$a. \quad x + y = (x_1 + y_1, \dots, x_n + y_n) = y + x.$$

$$b. \quad (x + y) + z = (x_1 + y_1 + z_1, \dots, x_n + y_n + z_n) = x + (y + z).$$

c. Pour le vecteur $0 = (0, 0, \dots, 0)$ (« vecteur nul »), on a l'égalité $x + 0 = (x_1 + 0, \dots, x_n + 0) = x$.

d. Pour tout $x \in R_n$, il existe un élément *opposé* y tel que $x + y = 0$; notamment, on peut poser

$$y = (-x_1, \dots, -x_n).$$

En même temps que les axiomes a-d, sont valables toutes leurs conséquences exprimées dans les propositions du paragraphe 1.3.

2.63. Définissons dans l'espace R_n l'opération de multiplication par un nombre arbitraire $\alpha \in R$ par la formule

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n). \quad (1)$$

Les formules suivantes sont, évidemment, justes :

$$\alpha(x + y) = \alpha x + \alpha y; \quad (2)$$

$$(\alpha + \beta)x = \alpha x + \beta x; \quad (3)$$

$$1 \cdot x = x; \quad (4)$$

$$\alpha(\beta x) = (\alpha\beta)x. \quad (5)$$

2.64. Les vecteurs

$$x^{(1)} = (x_1^{(1)}, \dots, x_n^{(1)}), \dots, x^{(m)} = (x_1^{(m)}, \dots, x_n^{(m)})$$

de l'espace R_n sont dits *linéairement dépendants* s'il existe des constantes $C_1, \dots, C_m$, pas toutes nulles, telles que

$$C_1 x^{(1)} + \dots + C_m x^{(m)} = 0. \quad (1)$$

Et si l'égalité (1) n'est possible que pour $C_1 = \dots = C_m = 0$, les vecteurs $x^{(1)}, \dots, x^{(m)}$ sont dits *linéairement indépendants*. Etant donnés 2.62 (1) et 2.63 (1), l'égalité vectorielle (1) équivaut à n égalités numériques

$$C_1 x_1^{(1)} + \dots + C_m x_1^{(m)} = 0,$$

$$C_1 x_n^{(1)} + \dots + C_m x_n^{(m)} = 0.$$

C'est pourquoi, en vertu des théorèmes connus de la théorie des systèmes d'équations linéaires *), la dépendance linéaire des vecteurs $x^{(1)}, \dots, x^{(m)}$ est équivalente au fait que le rang de la matrice $\|x_j^{(i)}\|$ est inférieur à m . En particulier, toute matrice carrée $\|x_j^{(i)}\|$ de n lignes et colonnes et dont le déterminant est non nul détermine dans R_n un système de n vecteurs linéairement indépendants. Et n'importe quels $n + 1$ vecteurs de l'espace R_n sont linéairement dépendants.

2.65. Soit $f_1, \dots, f_n$ un système de n vecteurs linéairement indépendants de l'espace R_n . Alors, pour tout vecteur $g \in R_n$, les

*) Voir, par exemple, [13].

vecteurs $f_1, \dots, f_n$, g étant linéairement dépendants, il existe la relation

$$C_0 g + C_1 f_1 + \dots + C_n f_n = 0, \quad (1)$$

où nécessairement $C_0 \neq 0$, sinon les vecteurs $f_1, \dots, f_n$ seraient linéairement dépendants. Donc, l'égalité (1) peut être résolue par rapport à g ; nous obtenons l'égalité de la forme

$$g = \alpha_1 f_1 + \dots + \alpha_n f_n. \quad (2)$$

Les nombres $\alpha_1, \dots, \alpha_n$ s'appellent *coordonnées* du vecteur g par rapport à la base $f_1, \dots, f_n$.

Ainsi, en choisissant pour vecteurs $f_1, \dots, f_n$ les vecteurs

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 1), \quad (3)$$

on obtient pour tout vecteur $x = (x_1, \dots, x_n) \in R_n$:

$$x = x_1 e_1 + \dots + x_n e_n,$$

de sorte que les nombres $x_1, \dots, x_n$ sont les coordonnées du vecteur x par rapport à la base spéciale (dite canonique) $e_1, \dots, e_n$ (3).

2.66. Soit $f_1, \dots, f_n$ une base fixe dans l'espace R_n . Si $g = \alpha_1 f_1 + \dots + \alpha_n f_n$ et $h = \beta_1 f_1 + \dots + \beta_n f_n$, alors on a d'après les propriétés 2.63 (2)-(3):

$$g + h = (\alpha_1 + \beta_1) f_1 + \dots + (\alpha_n + \beta_n) f_n,$$

$$Cg = C\alpha_1 f_1 + \dots + C\alpha_n f_n \text{ pour tout nombre } C,$$

de sorte que, à l'addition des vecteurs et à la multiplication d'un vecteur par un nombre C , il correspond, dans n'importe quelle base, l'addition des coordonnées respectives et leur multiplication par ce nombre C . On peut dire que l'application qui à tout vecteur $x = (x_1, \dots, x_n)$ fait correspondre un vecteur $\alpha = (\alpha_1, \dots, \alpha_n)$, où $\alpha_1, \dots, \alpha_n$ sont les coordonnées du vecteur x par rapport à une base fixe $f_1, \dots, f_n$, est un automorphisme (2.52) de l'espace n -dimensionnel R_n .

2.67. D'une façon générale, en accord avec la définition de l'automorphisme d'une structure mathématique (2.52), nous appelons *automorphisme de l'espace n -dimensionnel R_n* une application bijective $x' = A(x)$ de l'espace R_n sur lui-même telle que les conditions suivantes sont remplies:

a. $A(\alpha x) = \alpha A(x)$ pour tout $x \in R_n$ et tout $\alpha \in R$.

b. $A(x + y) = A(x) + A(y)$ quels que soient x et y de R_n .

Il est aisé de tirer des propriétés a et b une formule plus générale:

c. $A\left(\sum_{k=1}^p \alpha_k x_k\right) = \sum_{k=1}^p \alpha_k A(x_k)$ quels que soient les vecteurs

$x_1, \dots, x_p$ et $\alpha_1, \dots, \alpha_p$ réels.

(éléments omis sont nuls) multiplie par λ_j la j -ième coordonnée de chaque vecteur. Il s'appelle dilatation de λ_j fois suivant le j -ième axe. L'automorphisme défini par la matrice

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix} \quad (\lambda_1 \lambda_2 \dots \lambda_n \neq 0),$$

effectue une dilatation suivant chaque axe: il multiplie la première coordonnée par λ_1 , la deuxième par λ_2 , etc. Si tous les nombres $\lambda_1, \dots, \lambda_n$ se confondent, alors l'automorphisme correspondant défini par la matrice

$$\begin{pmatrix} \lambda & & 0 \\ & \lambda & \\ & & \ddots \\ 0 & & & \lambda \end{pmatrix} \quad (\lambda \neq 0),$$

transforme tout vecteur x en le vecteur λx . Un tel automorphisme s'appelle *homothétie de rapport λ* .

2.68. Espace euclidien n -dimensionnel. a. Soient x et y deux vecteurs quelconques d'un espace n -dimensionnel R_n : $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$. Le nombre

$$(x, y) = \sum_{k=1}^n \xi_k \eta_k \quad (1)$$

s'appelle *produit scalaire* des vecteurs x et y . Le nombre

$$|x| = \sqrt{(x, x)} = \sqrt{\sum_{k=1}^n \xi_k^2} > 0$$

s'appelle *longueur* ou *norme* du vecteur x . Un vecteur x pour lequel $|x| = 1$ est dit *unitaire* ou *normé*. Nous verrons plus bas que le produit scalaire permet de construire dans l'espace R_n une géométrie utilisant les notions de longueur, d'angle et, plus tard, d'aire et de volume. L'espace R_n muni du produit scalaire (1) est appelé *espace euclidien n -dimensionnel* et désigné par E_n .

b. Il est aisé de prouver que le produit scalaire (1) possède les propriétés suivantes:

- 1) $(y, x) = (x, y)$ quels que soient deux vecteurs x et y ;
- 2) $(\alpha x, y) = \alpha (x, y)$ quels que soient x et y de E_n et α réel;
- 3) $(x + y, z) = (x, z) + (y, z)$ pour n'importe quels x, y, z de E_n ;

4) $(x, x) \geq 0$ pour tout $x \in E_n$, de sorte que $(x, x) = 0$ dans le seul cas $x = 0$.

En partant des propriétés 1)-2)-3) on peut déduire la formule générale

$$\left(\sum_{i=1}^k \alpha_i x_i, \sum_{j=1}^m \beta_j y_j \right) = \sum_{i=1}^k \sum_{j=1}^m \alpha_i \beta_j (x_i, y_j), \quad (2)$$

qui est valable quels que soient des vecteurs $x_1, \dots, x_k, y_1, \dots, y_m$ et des nombres réels $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_m$.

e. Établissons l'inégalité de Cauchy :

$$|(x, y)| \leq |x| \cdot |y|. \quad (3)$$

La démonstration est basée sur les propriétés 1)-2)-3) et la première partie de 4). Considérons l'expression

$$\varphi(\lambda) = (x - \lambda y, x - \lambda y) = (x, x) - 2\lambda(x, y) + \lambda^2(y, y) = A - 2\lambda B + \lambda^2 C^2,$$

où $A = (x, x)$, $B = (x, y)$, $C = (y, y)$. D'après 4) on a $\varphi(\lambda) \geq 0$ pour tout $\lambda \in R$. Ainsi, le trinôme $\varphi(\lambda)$ ne peut avoir de racines réelles distinctes, sinon il prendrait des valeurs de signes différents. Il en résulte que la quantité $B^2 - AC = (x, y)^2 - (x, x)(y, y)$ (qui est sous le radical dans la formule des racines d'une équation du second degré) ne peut être positive. Par conséquent,

$$(x, y)^2 \leq (x, x)(y, y) = |x|^2 |y|^2,$$

d'où (3).

En coordonnées $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$, l'inégalité (3) s'écrit comme suit :

$$\left| \sum_{k=1}^n \xi_k \eta_k \right| \leq \sqrt{\sum_{k=1}^n \xi_k^2} \cdot \sqrt{\sum_{k=1}^n \eta_k^2} \quad (4)$$

Soulignons que, ainsi écrite, elle est valable pour n'importe quels nombres réels $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n$.

d. Vérifions l'inégalité de triangle

$$|x + y| \leq |x| + |y|. \quad (5)$$

En effet, quels que soient deux vecteurs x et y , on a

$$\begin{aligned} |x + y|^2 &= (x + y, x + y) = (x, x) + 2(x, y) + (y, y) \leq |x|^2 + \\ &+ 2|x||y| + |y|^2 = (|x| + |y|)^2, \end{aligned}$$

d'où (5). En coordonnées $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$, l'inégalité (5) prend la forme

$$\sqrt{\sum_{j=1}^n (\xi_j + \eta_j)^2} \leq \sqrt{\sum_{j=1}^n \xi_j^2} + \sqrt{\sum_{j=1}^n \eta_j^2};$$

tout comme (4), elle est valable pour n'importe quels nombres réels $\xi_1, \dots, \xi_n$ et $\eta_1, \dots, \eta_n$.

On peut également démontrer, pour $y = 0$, $z = (\zeta_1, \dots, \zeta_n)$, l'inégalité

$$\left| \sqrt{\sum_1^n \xi_k^2} - \sqrt{\sum_1^n \zeta_k^2} \right| \leq \sqrt{\sum_1^n (\xi_k - \zeta_k)^2} \quad (\text{cf. 3.11 (2)}). \quad (6)$$

Les inégalités (4) et (6) qui sont valables pour $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n, \zeta_1, \dots, \zeta_n$ réels quelconques sont très souvent appliquées en analyse aux estimations de toutes sortes, même sans aucun rapport avec l'origine géométrique de ces inégalités, la théorie des espaces métriques.

Signalons encore une chaîne utile d'inégalités:

$$\begin{aligned} \max_{k \leq n} |\xi_k - \eta_k| &\leq \rho(x, y) = \sqrt{\sum_1^n (\xi_k - \eta_k)^2} \leq \\ &\leq \sqrt{n} \cdot \max_{1 \leq k \leq n} |\xi_k - \eta_k|. \end{aligned} \quad (7)$$

Pour l'établir, passons de l'inégalité évidente

$$(\xi_k - \eta_k)^2 \leq \sum_1^n (\xi_k - \eta_k)^2$$

à sa conséquence (1.55.f)

$$|\xi_k - \eta_k| \leq \sqrt{\sum_1^n (\xi_k - \eta_k)^2}.$$

En passant au maximum sur k dans le premier membre on obtient la première inégalité (7). D'autre part, on a

$$\begin{aligned} \sum_1^n (\xi_k - \eta_k)^2 &\leq \sum_1^n \max_k (\xi_k - \eta_k)^2 = n \cdot \max_k (\xi_k - \eta_k)^2 = \\ &= n [\max_k |\xi_k - \eta_k|]^2, \end{aligned}$$

d'où la deuxième inégalité (7).

e. La définition générale de l'angle formé par deux vecteurs sera donné plus bas (5.74). Mais nous pouvons déjà formuler la définition d'orthogonalité. Deux vecteurs x et y sont dits *orthogonaux* si $(x, y) = 0$. En particulier, le vecteur nul $(0, \dots, 0)$ est orthogonal à tout vecteur $x \in E_n$. Si deux vecteurs x et y sont orthogonaux, alors

$$|x + y|^2 = (x + y, x + y) = (x, x) + 2(x, y) + (y, y) = |x|^2 + |y|^2;$$

c'est le théorème classique de Pythagore.

Des vecteurs $x_1, \dots, x_m$ non nuls et deux à deux orthogonaux sont linéairement indépendants; en effet, si l'on multiplie scalairement l'égalité

$$C_1 x_1 + \dots + C_m x_m = 0$$

par un vecteur x_k , on aura, vu (2) et l'hypothèse d'orthogonalité,

$$C_k(x_k, x_k) = 0.$$

Comme $(x_k, x_k) \neq 0$, il en résulte $C_k = 0$, ce qu'il nous fallait.

Notons en conclusion que nous avons construit l'espace vectoriel réel n -dimensionnel R_n et l'espace euclidien n -dimensionnel E_n en partant de la structure déjà connue de nombres réels. Dans ce qui suit (chapitre 11), nous verrons que les espaces vectoriels, et les espaces euclidiens en particulier, peuvent être traités eux-mêmes en tant que structures mathématiques, avec leurs axiomes. On a vu (2.53) que n'importe quelles deux structures de nombres réels sont isomorphes. D'une façon analogue, tout espace vectoriel n -dimensionnel défini axiomatiquement s'avère isomorphe à l'espace R_n , et tout espace euclidien n -dimensionnel défini axiomatiquement à l'espace E_n .

§ 2.7. Nombres complexes

2.71. Une question se pose tout naturellement: est-ce qu'il est possible d'introduire dans R_n une opération de multiplication des vecteurs de façon que le résultat soit encore un vecteur de R_n et que les axiomes de multiplication 1.22.a-e soient remplis? Il s'avère que c'est possible pour $n = 2$, par exemple, d'après la règle suivante. Soient $e_1 = (1, 0)$, $e_2 = (0, 1)$; tout vecteur $z \in R_2$ peut être mis sous la forme $z = (x, y) = xe_1 + ye_2$, où x et y sont les coordonnées du vecteur z par rapport à la base e_1, e_2 . Posons

$$e_1^2 = e_1, \quad e_1 e_2 = e_2 e_1 = e_2, \quad e_2^2 = -e_1.$$

Étendons cette multiplication sur tous les vecteurs $z = xe_1 + ye_2 \in R_2$ suivant la loi linéaire: si $w = ue_1 + ve_2$, alors

$$\begin{aligned} zw &= (xe_1 + ye_2)(ue_1 + ve_2) = xue_1^2 + xve_1e_2 + \\ &+ yue_2e_1 + yve_2^2 = (xu - yv)e_1 + (xv + yu)e_2. \end{aligned}$$

Bref, la définition de la multiplication dans R_2 est:

$$zw = (x, y)(u, v) = (xu - yv, xv + yu). \quad (1)$$

Evidemment, on a:

$$a. wz = (ux - vy, vx + uy) = zw$$

$$b. \text{ Si } \gamma = (\alpha, \beta), \text{ alors}$$

$$\begin{aligned} (zw)\gamma &= (xua - yva - xv\beta - uy\beta, xu\beta - yv\beta + xv\alpha + yu\alpha) = \\ &= (x(ua - v\beta) - y(va + u\beta), x(v\alpha + u\beta) + \\ &+ y(ua - v\beta)) = z(w\gamma) \end{aligned}$$

$$c. \text{ Pour le point } e = (1, 0) \text{ et pour tout } z = (u, v), \text{ on a}$$

$$(1, 0)(u, v) = (u, v)$$

d. Pour tout $(x, y) \neq (0, 0)$, il existe un élément inverse, notamment :

$$(u, v) = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right).$$

En effet,

$$(x, y)(u, v) = \left(\frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}, \frac{xy}{x^2 + y^2} - \frac{xy}{x^2 + y^2} \right) = (1, 0).$$

e. Pour n'importe quels $z = (x, y)$, $w = (u, v)$, $\gamma = (\alpha, \beta)$ on a :

$$(z + w)\gamma = z\gamma + w\gamma.$$

En effet,

$$\begin{aligned} (z + w)\gamma &= (x + u, y + v)(\alpha, \beta) = ((x + u)\alpha - \\ &\quad - (y + v)\beta, (x + u)\beta + (y + v)\alpha) = ((x\alpha - y\beta) + \\ &\quad + (u\alpha - v\beta), (x\beta + y\alpha) + (u\beta + v\alpha)) = (x\alpha - y\beta, x\beta + y\alpha) + \\ &\quad + (u\alpha - v\beta, u\beta + v\alpha) = (x, y)(\alpha, \beta) + (u, v)(\alpha, \beta) = z\gamma + w\gamma. \end{aligned}$$

Ainsi, pour la multiplication que nous avons introduite dans R_2 , tous les axiomes 1.22.a-e sont vérifiés. Par conséquent, toutes leurs conséquences signalées dans les propositions du § 1.4 ont lieu.

2.72. **D é f i n i t i o n.** L'espace R_2 muni des opérations d'addition 2.62 (1) et de multiplication 2.71 (1) est appelé *corps des nombres complexes* et désigné par C . L'ensemble R des nombres réels est immergé d'une façon biunivoque dans C d'après la règle

$$x \leftrightarrow (x, 0),$$

de sorte que les opérations d'addition et de multiplication sont conservées puisque

$$(x, 0) + (y, 0) = (x + y, 0), \quad (x, 0)(y, 0) = (xy, 0).$$

Pour cette raison, les nombres complexes $(x, 0)$ sont appelés nombres réels et notés $(x, 0) = x$.

Ainsi, le vecteur de base $e_1 = (1, 0)$ est désigné par 1 et le second vecteur de base $e_2 = (0, 1)$ par i . Tout nombre complexe (x, y) peut à présent être mis sous la forme $x + iy$. Par définition, nous avons $i^2 = e_2^2 = -e_1 = (-1, 0) = -1$, de sorte que, dans le domaine des nombres complexes, on peut extraire une racine carrée du nombre -1 . La seconde racine carrée de -1 est le nombre $-i$ puisque $(-i)^2 = (-1)^2 i^2 = -1$. Il n'existe pas d'autres racines carrées complexes de -1 : en effet, si $z^2 = -1$ pour un $z \in C$, alors $z^2 + 1 = (z + i)(z - i) = 0$, d'où, en vertu de 1.47.b (et compte tenu de 2.71), $z + i = 0$ ou bien $z - i = 0$, c'est-à-dire que $z = i$ ou bien $z = -i$.

La règle de multiplication 2.71 (1) peut être mise sous la forme :

$$(x + iy)(u + iv) = (xu - yv) + i(xv + yu),$$

ce qui correspond à la règle habituelle de la multiplication des binômes, compte tenu de l'égalité $i^2 = -1$.

Pour un $z = x + iy$ donné, le nombre x s'appelle *partie réelle* de z , et le nombre y sa *partie imaginaire*; nous notons $x = \operatorname{Re} z$, $y = \operatorname{Im} z$.

2.73. Deux nombres complexes $z = x + iy$ et $\bar{z} = x - iy$ sont dits *conjugués*.

Les propositions suivantes ont lieu :

- a. La relation $z = \bar{z}$ est équivalente au fait que z est réel.
 b. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
 c. $\overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2$ } quels que soient z_1 et z_2 complexes.

Démonstration. (a) La relation $x + iy = x - iy$ est équivalente à $y = 0$, c'est-à-dire que $z = x + iy$ est réel.

(b) Soient $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, on a

$$\overline{z_1 + z_2} = (x_1 + x_2) - i(y_1 + y_2) = (x_1 - iy_1) + (x_2 - iy_2) = \bar{z}_1 + \bar{z}_2.$$

(c) D'une manière analogue

$$\begin{aligned}\overline{z_1 z_2} &= (x_1 x_2 - y_1 y_2) - i(x_1 y_2 + x_2 y_1) = \\ &= (x_1 - iy_1)(x_2 - iy_2) = \bar{z}_1 \cdot \bar{z}_2.\end{aligned}$$

En particulier, (c) a pour conséquence la relation

d. $\overline{\alpha z} = \alpha \bar{z}$ pour tout α réel.

2.74. Unicité du système des nombres complexes.

Théorème. Si l'on introduit, dans l'espace bidimensionnel R_2 , une multiplication d'éléments qui vérifie les axiomes 1.22, alors il est possible de trouver dans R_2 deux vecteurs linéairement indépendants g_1 et g_2 tels que $g_1^2 = g_1$, $g_2^2 = -g_1$, $g_1 g_2 = g_2 g_1 = g_2$.

Démonstration. D'après l'axiome 1.22.c, il existe dans R_2 l'élément 1 différent de 0; posons $g_1 = 1$. Soit ensuite f_1 un vecteur quelconque linéairement indépendant de g_1 . Nous avons $f_1^2 = \alpha g_1 + \beta f_1$, avec α et β réels. Pour un vecteur $f_2 = f_1 + \gamma g_1 \neq 0$ (γ réel) nous avons $(f_1 + \gamma g_1)^2 = f_1^2 + 2\gamma f_1 + \gamma^2 g_1 = \alpha g_1 + \beta f_1 + 2\gamma f_1 + \gamma^2 g_1$, et si l'on pose $\gamma = -\beta/2$, on aura $f_2^2 = (\alpha + \gamma^2) g_1$. Le nombre réel $\alpha + \gamma^2$ ne peut être nul d'après 1.47.a. Il ne peut être positif non plus, sinon on aurait

$$(f_2 - \sqrt{\alpha + \gamma^2} g_1)(f_2 + \sqrt{\alpha + \gamma^2} g_1) = 0$$

en contradiction avec 1.47.h. Donc, $\alpha + \gamma^2 = -\delta$, $\delta > 0$. En remplaçant f_2 par $g_2 = \frac{1}{\sqrt{\delta}} f_2$, on trouve $g_2^2 = \frac{(\alpha + \gamma^2) g_1}{\delta} = -g_1$.

Les égalités $g_1 g_2 = g_2 g_1 = g_2$ découlent du fait que g_1 est l'unité. Le théorème est démontré.

Il résulte de ce théorème que l'on peut établir, entre n'importe quels deux espaces bidimensionnels munis chacun d'une multiplication vérifiant les axiomes 1.22, une correspondance biunivoque conservant l'opération de multiplication, c'est-à-dire que les structures en question sont isomorphes.

2.75. Automorphismes du système des nombres complexes. Nous avons vu dans 2.59 que le seul automorphisme du système des nombres réels est l'automorphisme identique $\xi \rightarrow \xi$. Voici la définition d'un automorphisme du domaine des nombres complexes: une application bijective $z \rightarrow z'$ du plan C sur lui-même s'appelle *automorphisme* si, pour n'importe quels z_1 et z_2 de C et pour tout α réel, les conditions suivantes sont satisfaites:

$$1) (z_1 + z_2)' = z_1' + z_2';$$

$$2) (z_1 z_2)' = z_1' z_2';$$

$$3) (\alpha z_1)' = \alpha z_1'.$$

Les propriétés 1) et 3) montrent que l'automorphisme considéré est un automorphisme du système C en tant qu'espace bidimensionnel (2.67). La propriété 2) dit que cet automorphisme doit conserver l'opération de multiplication.

Un automorphisme non identique du système C peut être donné par la règle

$$x + iy \rightarrow x - iy. \quad (1)$$

Montrons que le système C n'admet aucun autre automorphisme non identique. Soit $(x + iy)'$ l'élément en lequel se transforme un élément $x + iy$ par un automorphisme. Il résulte de 3) que

$$1' z' = (1 \cdot z)' = z',$$

de sorte que $1'$ est une unité du système C ; en vertu de l'unicité de l'unité, on a $1' = 1$. D'après 2), pour tout x réel, $x' = (x \cdot 1)' = x \cdot 1' = x \cdot 1 = x$, de sorte que l'automorphisme conserve tout l'ensemble des nombres réels. Ensuite, d'après 2), on a:

$$(i')^2 = i^2 = -1,$$

d'où i' est soit i , soit $-i$ (2.72). Dans le premier cas, on trouve

$$(x + iy)' = x' + i'y' = x + iy$$

et dans le second

$$(x + iy)' = x' + i'y' = x - iy,$$

ce qui démontre notre proposition.

L'étude des nombres complexes sera continuée dans le chapitre 5 (5.72 et suivants).

2.76. Est-il possible, pour $n > 2$, d'introduire une opération de multiplication des vecteurs de l'espace R_n vérifiant les axiomes 1.22.a-e? Un théorème de Frobenius y donne une réponse négative. Nous ne nous arrêtons pas sur la démonstration de ce théorème intéressant [1].

§ 2.8. Notion générale de fonction. Graphique

2.81. Soient X et Y des ensembles constitués d'éléments x et y respectivement. Supposons qu'on fasse correspondre de façon f à tout élément $x \in X$ un élément $y \in Y$; alors la correspondance $x \rightarrow y$ (ou $y = f(x)$) s'appelle *fonction ayant X pour domaine de définition et Y pour domaine de valeurs*.

Dans ce contexte, x s'appelle *variable indépendante* (ou *argument*) de la fonction $f(x)$ et y *variable dépendante* (ou *valeur*) de la fonction.

Soulignons que la définition d'une fonction n'exige pas que tout $y \in Y$ soit une valeur $f(x)$ pour un certain $x \in X$ ni qu'à des valeurs différentes de x correspondent des valeurs différentes de y . (Si ces deux conditions sont vérifiées, on a affaire à une correspondance biunivoque (2.21) qui est un cas spécial d'une fonction.) Si le domaine de valeurs Y d'une fonction $f(x)$ est la droite numérique R , alors la fonction $f(x)$ s'appelle *fonction numérique*. Si Y est l'espace vectoriel R_n (§ 2.6), la fonction $f(x)$ est une *fonction vectorielle*.

Si le domaine de définition X d'une fonction $f(x)$ est la droite numérique R , ou la droite numérique achevée $\bar{R}$ (1.91), ou bien un ensemble $E \subset \bar{R}$, alors $f(x)$ s'appelle *fonction d'une variable réelle*.

Si le domaine de définition d'une fonction $f(x)$ est l'ensemble des nombres naturels $1, 2, 3, \dots$, alors la fonction $f(x)$ désignée autrement par $f(n)$ ou bien par f_n est appelée *suite* de points de l'ensemble Y . Signalons que la notion de suite de points d'un ensemble ne se réduit pas à la notion de sous-ensemble (au plus dénombrable): une suite est organiquement liée à la numérotation; un ensemble ne change pas en permutant ses éléments, tandis qu'en faisant de même avec une suite on aboutit à une autre suite.

2.82. Introduisons la notion de produit cartésien de deux ensembles quelconques. On appelle *produit cartésien* de deux ensembles X et Y l'ensemble $P(X, Y)$ de tous les couples (x, y) , où le premier élément appartient à l'ensemble X et le deuxième à l'ensemble Y . L'égalité de deux couples (x, y) et (x', y') est définie par les conditions $x' = x, y' = y$. Ainsi, le plan réel (R_2) est le produit cartésien de deux droites réelles.

Fixons, dans l'ensemble Y , un élément y_0 et considérons le sous-ensemble de tous les points $(x, y_0) \in P(X, Y)$. Ce sous-ensemble qui est évidemment équipotent à l'ensemble X s'appelle *coupe* de $P(X, Y)$ suivant l'élément y_0 . Le produit cartésien $P(X, Y)$ est la réunion des coupes différentes équipotentes à l'ensemble X (et les unes aux autres).

2.83. S'il y a une fonction $y = f(x)$ ayant X pour domaine de définition et Y pour domaine de valeurs, nous appelons son *graphique* le sous-ensemble du produit cartésien $P(X, Y)$, qui est formé des

couples (x, y) pour lesquels $y = f(x)$. En cas de $X = R_1$ et $Y = R_1$, cette définition coïncide avec la définition ordinaire du graphique d'une fonction réelle de la variable numérique. Dans des cas pratiquement importants, le graphique d'une fonction numérique d'une variable réelle représente une ligne du plan (x, y) . Si le domaine de définition X d'une fonction $f(x)$ est le plan $x = (x_1, x_2)$, et le domaine de valeurs la droite numérique, alors son graphique est un ensemble dans R_3 que l'on peut souvent se représenter sous la forme d'une surface. Si le domaine de définition X d'une fonction $f(x)$ est l'axe réel et le domaine de valeurs Y le plan $y = (y_1, y_2)$, alors le graphique de la fonction $f(x)$ est toujours un ensemble dans R_3 ; mais, cette fois, il est mieux de se le représenter sous la forme d'une ligne qui coupe tout plan $x = \text{const}$ en un seul point. Tous les exemples considérés représentent autant d'objets des plus importants de l'analyse et seront étudiés dans la suite d'une façon systématique.

2.84. Remarque sur les fonctions « univoques » et « multivoques ». En vertu de la définition même 2.81, toute fonction fait correspondre à tout élément $x \in X$ un élément unique $y \in Y$ et, en ce sens, elle est « univoque ». On dit parfois « fonction multivoque » et on sous-entend qu'à tout élément $x \in X$ correspond non pas un seul mais plusieurs éléments de l'ensemble Y . On pourrait étendre la définition d'une fonction dans ce sens. Pourtant, les difficultés qui surgiraient lorsqu'on aurait à définir les opérations sur de telles « fonctions » nous font abandonner cette idée. Tout de même, des fois, le terme « fonction multivoque » s'avérera utile. Notamment, on réunit parfois plusieurs fonctions univoques en une formule (pour abréger la notation, par exemple); une telle réunion est appelée fonction « multivoque ». Par exemple, la « fonction bivoque »

$$y = x \pm \sqrt{1 - x^2} \quad (-1 \leq x \leq 1)$$

est tout simplement la réunion de deux fonctions univoques

$$y_1 = x + \sqrt{1 - x^2}, \quad y_2 = x - \sqrt{1 - x^2} \quad (-1 \leq x \leq 1).$$

Exercices

1. Démontrer que les ensembles ci-dessous sont dénombrables:

- Ensemble de tous les intervalles $a < x < b$ aux extrémités rationnelles.
- Ensemble de tous les polygones plans à un nombre fini de côtés et de sommets aux points rationnels.

2. Démontrer que chacun des ensembles suivants est ou bien fini ou bien dénombrable:

- Ensemble d'intervalles ouverts disjoints deux à deux donnés sur un axe.
- Ensemble de courbes fermées disjointes deux à deux (données dans le plan) chacune desquelles se coupe comme un huit.

c) Ensemble M des nombres réels positifs tels que toutes les sommes finies

$$\sum x_j, x_j \in M, \text{ soient majorées par un nombre fixe } A.$$

R o m a r q u e. V. Grouhino et V. Palamodov ont démontré les propositions analogues pour tout ensemble de figures disjointes du plan qui possèdent des points triples (comme la lettre T) et aussi pour tout ensemble de figures disjointes dans l'espace qui possèdent ou bien des points singuliers du type « punaise », ou bien des parties du type « feuille de Möbius » *).

3. Décomposer l'ensemble des nombres naturels $1, 2, \dots$ en une réunion dénombrable d'ensembles dénombrables disjoints deux à deux.

4. (Problème amusant). I. Un jour le mathématicien X reçoit ses amis, les frères N . Ceux-ci enlèvent leurs chapeaux et les laissent dans l'antichambre. Au moment de partir on s'aperçoit, à la grande confusion de l'hôte, qu'il manque un chapeau. Or, personne n'est entrée dans l'antichambre durant la visite.

II. Une autre fois les frères N venus voir X accrochent encore leurs chapeaux sur le porte-chapeaux. Quand ils s'apprêtent à partir, ils trouvent un chapeau de trop. Pourtant l'hôte et les visiteurs savent bien qu'avant leur arrivée le porte-chapeaux était vide.

III. La fois suivante, les frères mettent leurs chapeaux et partent. L'hôte les accompagne dans la rue, puis rentre et voit sur le porte-chapeaux autant de chapeaux qu'il y en a eu avant le départ des visiteurs.

IV. Enfin, la quatrième fois, les frères viennent tête nue et en partant prennent des chapeaux qui sont restés de la dernière visite. En rentrant après avoir accompagné ses visiteurs l'hôte trouve autant de chapeaux qu'il y en a eu avant l'arrivée des frères.

Comment expliquer ces faits paradoxaux?

5. Si l'on ajoute un ensemble B fini ou dénombrable à un ensemble infini A , on obtient un ensemble équipotent à l'ensemble initial A .

6. Démontrer que l'ensemble i des nombres irrationnels et l'ensemble T des nombres transcendants ont chacun la puissance du continu.

7. L'ensemble de toutes les suites composées de zéro et d'unités a la puissance du continu.

8. L'ensemble de toutes les suites croissantes de nombres naturels

$$(0 <) k_1 < k_2 < \dots < k_n < \dots \quad (1)$$

a la puissance du continu.

9. L'ensemble de toutes les suites de nombres naturels (pas obligatoirement croissantes)

$$m_1, m_2, \dots, m_n, \dots \quad (2)$$

a la puissance du continu.

10. L'ensemble de toutes les suites de nombres réels

$$\xi = (\xi_1, \xi_2, \dots, \xi_n, \dots)$$

a la puissance du continu.

11. L'ensemble de tous les points de l'espace n -dimensionnel R_n , pour tout $n = 1, 2, \dots$, a la puissance du continu.

12. La puissance de l'ensemble E de toutes les fonctions $y = f(x)$ définies sur l'intervalle $[0, 1]$ et dont le domaine de valeurs ne comprend que deux points distincts n'est pas celle du continu.

13. L'ensemble de tous les sous-ensembles d'un ensemble A n'est pas équipotent à A .

Historique

Les idées principales de la théorie des ensembles, formulées pour la première fois à la fin du XIX^e siècle par Cantor, ont pénétré depuis dans les branches les plus diverses des mathématiques en parachevant, ou peu s'en faut, leur langage. (Voir F. Hausdorff, *Mengenlehre*, Berlin (de Gruyter), 1927). A la limite des

*) Voir [7].

XIX^e et XX^e siècles, on s'aperçoit que la logique traditionnelle est génératrice de contradictions dans la théorie des ensembles. Cette situation dramatique qui met en doute tous les résultats des mathématiques trouve sa résolution grâce à la création des systèmes axiomatiques de la théorie des ensembles qui fixent non seulement les propriétés des objets mais aussi les moyens logiques admissibles. (Voir P. Cohen, *Set theory and the continuum hypothesis*, N. Y.-Amst., Benjamin, 1966.)

Le rôle que jouent les structures dans les mathématiques est mis en relief par Bourbaki dans l'article « L'architecture des mathématiques » (1948). L'œuvre en plusieurs volumes de Bourbaki « Éléments de mathématique » (le premier fascicule paraît en 1939) est consacrée à la représentation de toutes les mathématiques sous la forme d'un système de structures. A l'heure actuelle l'édition est encore loin d'être terminée; les mathématiques étant impuissables, il paraît que l'œuvre ne sera jamais terminée d'une façon plus ou moins naturelle.

Les systèmes de coordonnées (dans le plan et dans l'espace tridimensionnel) sont introduits par Descartes (1637) dont la *Géométrie* exerce une influence immense sur le développement des mathématiques.

Les espaces à plusieurs dimensions sont introduits par Cayley et Graessmann (1846). Les nombres complexes apparaissent beaucoup plus tôt (XVI^e siècle) chez les algébristes italiens qui s'en servent en résolvant les équations algébriques, les règles d'opérations sur les expressions de la forme $a + b\sqrt{-1}$ sont déjà décrites dans l'« Algebra » de Bombelli (1579). Puisque, à l'époque, on ne pouvait attribuer aucun sens intuitif aux nombres complexes, on les appelait « imaginaires », « absurdes », etc., jusqu'à ce que Gauss (1797) et, indépendamment de lui, Wessel (1798) et Argand (1806) ne les interprètent comme points du plan. En 1903, Frobenius établit la non-existence, pour $n > 2$, d'un espace n -dimensionnel muni de l'addition et de la multiplication satisfaisant à tous les axiomes 1.21-1.22.

A la fin du XIX^e siècle, avec le développement de la théorie des ensembles, il devient clair que la représentation formelle du nombre complexe $a + b\sqrt{-1}$ par un couple (a, b) ne le cède en rien à la représentation géométrique.

La définition générale d'une fonction (2.81) est formulée (pour les fonctions numériques) par Lobatchevski (1834) et Dirichlet (1837) et avant, mais sous une forme moins nette, par Euler (1751); elle couronne le travail de longue haleine de mathématiciens du XVIII^e siècle; cette définition sépare la notion générale de fonction de la notion (qui s'élargit toujours avec le développement des mathématiques) de représentation analytique. Voir [2].

Espaces métriques

La seule chose qui manque actuellement à la théorie des ensembles pour occuper sa place dans l'analyse, c'est la conception générale de passage à limite.

J. Hadamard (1900)

Les espaces métriques représentent l'une des plus importantes structures mathématiques. En particulier, elle est à la base de la théorie générale des limites exposée dans le chapitre 4.

§ 3.1. Définitions et exemples

3.11. Définition. Un ensemble M d'éléments (« points ») $x, y, \dots$ s'appelle *espace métrique* s'il y a une règle qui permet de faire correspondre à deux points quelconques x, y un nombre $\rho(x, y)$ (« distance de x à y ») et satisfait aux conditions (axiomes) suivantes :

a. $\rho(x, y) > 0$ pour $x \neq y$; $\rho(x, x) = 0$ pour tout x .

b. $\rho(x, y) = \rho(y, x)$ quels que soient x et y (symétrie de la distance).

c. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ quels que soient x, y, z (« inégalité de triangle »).

L'inégalité de triangle a l'origine géométrique suivante : en géométrie élémentaire, la longueur du côté xz d'un triangle xyz ne dépasse pas la somme des longueurs de deux autres côtés xy et yz .

La règle qui permet, d'après deux points x, y de l'espace M , de trouver le nombre $\rho(x, y)$ est appelée *métrique* de l'espace M .

Notons que tout sous-ensemble $M' \subset M$ d'un espace métrique M est un espace métrique par rapport à la métrique de l'espace M .

Signalons quelques conséquences simples des axiomes a-c :

d. Pour des éléments quelconques $x_1, \dots, x_n$ de l'espace métrique M , on a l'inégalité

$$\rho(x_1, x_n) \leq \rho(x_1, x_2) + \rho(x_2, x_3) + \dots + \rho(x_{n-1}, x_n)$$

(une généralisation du théorème connu de la géométrie élémentaire : la longueur du segment joignant la première et la dernière extrémité d'une ligne polygonale ne dépasse pas la somme des longueurs des côtés de la ligne). Pour le démontrer, on applique successivement l'axiome c :

$$\begin{aligned} \rho(x_1, x_n) &\leq \rho(x_1, x_2) + \rho(x_2, x_n) \leq \\ &\leq \rho(x_1, x_2) + \rho(x_2, x_3) + \rho(x_3, x_n) \leq \\ &\leq \rho(x_1, x_2) + \rho(x_2, x_3) + \dots + \rho(x_{n-1}, x_n). \end{aligned}$$

e. Pour quatre points quelconques x, y, z, v de l'espace métrique M on a l'« inégalité de quadrilatère »

$$|\rho(x, y) - \rho(z, v)| \leq \rho(x, z) + \rho(y, v). \quad (1)$$

En effet, nous avons, d'après d :

$$\begin{aligned} \rho(x, y) &\leq \rho(x, z) + \rho(z, v) + \rho(v, y), \\ \rho(z, v) &\leq \rho(z, x) + \rho(x, y) + \rho(y, v) \end{aligned}$$

ou bien

$$\begin{aligned} \rho(x, y) - \rho(z, v) &\leq \rho(x, z) + \rho(v, y), \\ \rho(z, v) - \rho(x, y) &\leq \rho(z, x) + \rho(y, v). \end{aligned}$$

Les seconds membres de ces inégalités se confondent par l'axiome b. Les premiers sont opposés; il en résulte (1).

f. En posant $v = y$ dans l'inégalité (1) on obtient :

$$|\rho(x, y) - \rho(z, y)| \leq \rho(x, z) \quad (2)$$

(en géométrie élémentaire, la différence des longueurs de deux côtés d'un triangle ne dépasse pas la longueur du troisième).

3.12. a. Un ensemble B dans un espace métrique M est dit *borné* si les distances d'un point quelconque $a \in M$ à tous les points $b \in B$ sont majorées par une constante fixe. Dans ce cas, les distances de n'importe quel autre point fixe c à tous les points de l'ensemble B sont aussi majorées par une constante fixe, de même que les distances entre deux points quelconques de l'ensemble B . Tout cela résulte de l'inégalité de triangle

$$\rho(b, c) \leq \rho(b, a) + \rho(a, c).$$

La quantité

$$\text{diam } B = \sup_{x \in B, y \in B} \rho(x, y)$$

s'appelle *diamètre* de l'ensemble B (borné).

b. Dire que, pour $M = R_1$ (axe réel), un ensemble B est borné au sens de a équivaut à dire qu'il est borné au sens ordinaire (1.66),

puisque, dans les deux cas, cela signifie que B fait partie d'un intervalle (fini).

c. Un ensemble B dans un espace métrique M n'est pas borné si, pour tout C réel, il existe dans l'ensemble B deux points x, y avec $p(x, y) > C$. Dans ce cas, pour tout C et pour tout point fixe $a \in M$, on peut trouver un point $x \in B$ pour lequel $p(a, x) > C$; sinon l'ensemble B serait borné et, d'après a, les distances entre ses deux points quelconques seraient majorées.

d. Les définitions d'un espace métrique borné ou non borné s'obtiennent respectivement de a et c pour $B = M$.

e. L'ensemble de tous les points x d'un espace métrique M dont les distances à un point donné x_0 sont inférieures à un nombre donné $r > 0$, de sorte que

$$p(x, x_0) < r,$$

s'appelle *boule* (plus exactement, *boule ouverte*) de rayon r ; le point x_0 est le *centre* de cette boule. L'ensemble de tous les points x vérifiant l'inégalité

$$p(x, x_0) \leq r$$

s'appelle *boule fermée* de rayon r centrée en x_0 . Enfin, les points se trouvant à une distance fixe r d'un point x_0 , de sorte que

$$p(x, x_0) = r,$$

forment la *sphère* de rayon r centrée en x_0 .

f. Toute boule de centre x_0 s'appelle *voisinage* de ce point. Un point x_0 s'appelle *point intérieur* d'un ensemble $E \subset M$ s'il a un voisinage inclus dans E .

3.13. Exemples.

a. Tout ensemble M sur la droite réelle R , est un espace métrique avec pour distance $p(x, y) = |x - y|$.

La vérification des conditions 2.11. a-c résulte des propriétés connues du module (1.54.b). Si M est toute la droite, alors une boule ouverte de rayon r centrée en un point x_0 est l'intervalle

$$|x_0 - x| < r;$$

une boule fermée de rayon r centrée en un point x_0 est l'intervalle

$$|x - x_0| \leq r;$$

une sphère de rayon r centrée en un point x_0 est deux points

$$x = x_0 \pm r.$$

b. De même, un ensemble M dans le plan R_2 ou bien dans l'espace tridimensionnel R_3 est un espace métrique si l'on choisit pour distance entre les points (dans R_3 , pour fixer les idées) $x = (\xi_1, \xi_2, \xi_3)$

et $y = (\eta_1, \eta_2, \eta_3)$ la distance géométrique ordinaire:

$$\rho(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^2}.$$

L'inégalité de triangle (axiome c) se réduit ici à l'inégalité géométrique connue: un côté d'un triangle ne dépasse pas la somme de deux autres. Un exemple plus général est considéré dans le numéro suivant.

3.14.a. Considérons un espace euclidien n -dimensionnel E_n (2.68). Rappelons que tout vecteur $x = (\xi_1, \dots, \xi_n)$ de cet espace a sa

longueur $|x| = \sqrt{\sum_{k=1}^n \xi_k^2}$. Définissons à présent la distance entre deux éléments $x = (\xi_1, \dots, \xi_n)$ et $y = (\eta_1, \dots, \eta_n)$ de E_n par la formule:

$$\rho(x, y) = |x - y| = \sqrt{\sum_{k=1}^n (\xi_k - \eta_k)^2}. \quad (1)$$

Les axiomes de la distance 3.11.a-b sont ici immédiats. Pour vérifier l'axiome 3.11.c, remplaçons dans l'inégalité 2.68 (5) x par $x - y$, y par $y - z$, donc $x + y$ par $x - z$; nous aurons

$$\rho(x, z) = |x - z| \leq |x - y| + |y - z| = \rho(x, y) + \rho(y, z),$$

ce qu'il nous fallait. Ainsi, l'espace E_n muni de la distance (1), de même que tout sous-ensemble $E \subset E_n$ avec la même distance, est un espace métrique.

b. Comment décrire les ensembles bornés dans l'espace n -dimensionnel R_n ? Si un ensemble $A \subset R_n$ est borné, alors les distances de tous les points $x = (\xi_1, \dots, \xi_n) \in A$ à un point quelconque de l'espace R_n , par exemple, au vecteur nul, sont majorées par une constante fixe. Autrement dit, il existe une constante b telle que, pour tout $x \in A$:

$$|x| = \sqrt{\sum_{k=1}^n \xi_k^2} \leq b.$$

Mais alors, pour tout $k = 1, 2, \dots, n$ et pour $x \in A$, on a

$$|\xi_k| \leq b, \quad (2)$$

c'est-à-dire que toute coordonnée des points $x \in A$ est bornée sur l'axe réel. Réciproquement, si un ensemble $A \subset R_n$ est tel que toute coordonnée de ses points est bornée sur l'axe réel, de sorte qu'il existe une constante b vérifiant l'inégalité (2) pour tout $k = 1, \dots, n$ et pour tout $x \in A$, alors

$$|x| = \sqrt{\sum_{k=1}^n \xi_k^2} \leq \sqrt{\sum_{k=1}^n b^2} = b\sqrt{n}.$$

de sorte que l'ensemble A est borné dans l'espace R_n . Donc, un ensemble $A \subset R_n$ est borné si, et seulement si, l'ensemble des valeurs de toutes les coordonnées des points $x \in A$ est borné sur la droite numérique.

3.15. Pour les espaces métriques en tant que structures mathématiques, on peut introduire, d'après 2.52, la définition d'isomorphisme qui, dans ce cas, s'appelle *isométrie*. Deux espaces métriques sont dits *isométriques* si l'on peut établir entre les éléments de ces espaces une correspondance biunivoque qui conserve les distances.

Par exemple, deux segments égaux sur la droite numérique avec sa métrique naturelle sont isométriques, deux segments de longueurs différentes ne le sont pas. Toute figure du plan est isométrique à sa symétrique par rapport à une droite. Notons qu'une isométrie dans l'espace n -dimensionnel qui transforme la somme de deux vecteurs quelconques x et y en la somme de leurs images (isométrie linéaire) conserve aussi bien le produit scalaire, ce qui résulte de l'égalité

$$(x + y, x + y) = (x, x) + 2(x, y) + (y, y).$$

Voici deux exemples simples d'applications isométriques d'un espace euclidien n -dimensionnel R_n sur lui-même :

a. La symétrie par rapport au plan $x_n = 0$: tout vecteur $x = (\xi_1, \dots, \xi_n)$ est transformé en le vecteur $x' = (\xi_1, \dots, \xi_{n-1}, -\xi_n)$.

b. La translation d'un vecteur $\beta = (\beta_1, \dots, \beta_n)$: tout vecteur $x = (\xi_1, \dots, \xi_n)$ se transforme en le vecteur

$$x + \beta = (\xi_1 + \beta_1, \dots, \xi_n + \beta_n).$$

Notons que, pour $\beta \neq 0$, cette application n'est pas un automorphisme de l'espace R_n (2.67): elle ne conserve pas l'origine des coordonnées.

c. En général, s'il existe une transformation isométrique d'un espace n -dimensionnel R_n sur lui-même par laquelle un ensemble $G \subset R_n$ devient l'ensemble $F \subset R_n$, alors ces ensembles G et F sont dits *géométriquement égaux* conformément à la terminologie de la géométrie élémentaire (bien que cela ne soit pas conforme à la définition même de l'égalité qui, en géométrie élémentaire, n'est formulée que pour les figures d'un type spécial).

3.16. *Métrisation du produit cartésien d'espaces métriques*. Soient M_1 et M_2 deux espaces métriques; attribuons les indices 1 et 2 aux symboles désignant la métrique et les points des espaces respectifs. Soit $M = M_1 \times M_2$ le produit cartésien des ensembles M_1 et M_2 (2.82), c'est-à-dire l'ensemble de tous les couples $x = (x_1, x_2)$, $x_1 \in M_1$, $x_2 \in M_2$. Définissons une métrique dans l'ensemble M suivant la règle:

si $x = (x_1, x_2)$, $y = (y_1, y_2)$, alors

$$p(x, y) = p[(x_1, x_2), (y_1, y_2)] = \max\{p_1(x_1, y_1), p_2(x_2, y_2)\}. \quad (1)$$

Vérifions les axiomes 3.11. a-c. Si $x \neq y$, alors $x_1 \neq y_1$ ou $x_2 \neq y_2$, c'est-à-dire que $p_1(x_1, y_1) > 0$ ou $p_2(x_2, y_2) > 0$; dans les deux cas, par définition (1), on a aussi $p(x, y) > 0$. Et si $x = y$, alors $x_1 = y_1$, $x_2 = y_2$, d'où $p_1(x_1, y_1) = 0$, $p_2(x_2, y_2) = 0$, donc $p(x, y) = 0$. L'axiome a est ainsi vérifié. Ensuite, $p(y, x) = \max\{p_1(y_1, x_1), p_2(y_2, x_2)\} = \max\{p_1(x_1, y_1), p_2(x_2, y_2)\} = p(x, y)$, de sorte que l'axiome b est vérifié. Enfin, soit $z = (z_1, z_2)$; évaluons $p(x, z) = \max\{p_1(x_1, z_1), p_2(x_2, z_2)\}$. Soit, pour fixer les idées, $p(x, z) = p_1(x_1, z_1)$; on a alors

$$\begin{aligned} p(x, z) = p_1(x_1, z_1) &\leq p_1(x_1, y_1) + p_1(y_1, z_1) \leq \\ &\leq p(x, y) + p(y, z), \end{aligned}$$

ce qui démontre l'axiome c.

Notons que la métrique (1), pour toute coupe (x, z) avec un $z \in M_2$ fixe (2.82), est identique à la métrique de l'espace M_1 , de sorte que toute coupe (x, z) est non seulement équipotente mais aussi isométrique à l'espace M_1 . De plus, la formule (1) n'est pas l'unique façon d'introduire une métrique dans le produit cartésien des espaces M_1 et M_2 . On aurait pu se servir d'autres formules, par exemple

$$p(x, y) = p_1(x_1, y_1) + p_2(x_2, y_2) \quad (2)$$

ou bien

$$p(x, y) = \sqrt{p_1^2(x_1, y_1) + p_2^2(x_2, y_2)}. \quad (3)$$

Nous laissons au lecteur le soin de vérifier les axiomes de la métrique dans les cas (2) et (3).

§ 3.2. Ensembles ouverts

3.21. Un ensemble U dans un espace métrique M s'appelle *ensemble ouvert* ou *domaine* si tout point x_0 de l'ensemble U est son point intérieur, i. e. il existe une boule ouverte (dont le rayon peut dépendre de x_0) centrée en x_0 et contenue dans U .

Ainsi, une boule ouverte centrée en x_1

$$U = \{x: p(x, x_1) < r\}$$

est un ensemble ouvert. En effet, soit $x_0 \in U$, de sorte que $p(x_0, x_1) = \theta < r$. Considérons une boule U_0 centrée en x_0 et de rayon $r_0 < r - \theta$; démontrons que U_0 est une partie de la boule U . En effet,

pour tout $x \in U_0$, on a par l'inégalité de triangle

$$\rho(x, x_1) \leq \rho(x, x_0) + \rho(x_0, x_1) < r_0 + \theta < r,$$

ce qu'il fallait démontrer.

3.22. Théorème. *Toute réunion d'une famille d'ensembles ouverts et l'intersection d'un nombre fini d'ensembles ouverts sont des ensembles ouverts.*

Démonstration. Le fait qu'une réunion d'ensembles ouverts est ouverte résulte de la définition même 3.21. Étudions le cas d'intersection d'un nombre fini d'ensembles ouverts. Soit un point x_0 qui appartient aux ensembles ouverts $U_1, U_2, \dots, U_m$ dont le premier contient une boule de rayon r_1 (centrée en x_0), le deuxième une boule de rayon r_2 , etc.; alors la boule de centre x_0 et de rayon $\min(r_1, r_2, \dots, r_m)$ est incluse dans chacun des ensembles $U_1, U_2, \dots, U_m$ et, par conséquent, dans leur intersection.

Le raisonnement n'est plus valable pour une famille infinie d'ensembles ouverts, puisque le minimum (plus exactement, la borne inférieure) d'un ensemble infini de nombres positifs peut être nul. En effet, l'intersection d'une famille infinie des ensembles ouverts

$$U_n = \left\{ x : \rho(x, x_0) < \frac{1}{n} \right\} \quad (n = 1, 2, \dots)$$

contient seuls les points x pour lesquels $\rho(x, x_0) = 0$, donc, d'après l'axiome 3.11.a, ne contient que le point x_0 ; cette intersection n'est en général pas un ensemble ouvert.

3.23. Tout intervalle (α, β) (borné ou non) de l'axe $-\infty < x < \infty$ est, évidemment, un ensemble ouvert. Une réunion finie ou dénombrable d'intervalles (α_v, β_v) ($v = 1, 2, \dots$) est aussi un ensemble ouvert. Montrons que *tout ensemble ouvert U de l'axe est une réunion finie ou dénombrable d'intervalles ouverts disjoints.*

Soit un point quelconque $x \in U$. Par définition, le point x appartient à l'ensemble U avec une boule, c'est-à-dire avec un intervalle ouvert de l'axe contenant le point x . Nous allons construire le plus grand intervalle ouvert contenant le point x et contenu entièrement dans l'ensemble U .

Désignons par S l'ensemble des points à droite de x et n'appartenant pas à U . Si S est vide, alors toute la demi-droite (x, ∞) est incluse dans U . Si l'ensemble S n'est pas vide, il existe bien sa borne inférieure ξ . Ce point ξ n'appartient évidemment pas à U , parce que tout point de l'ensemble U a un voisinage faisant partie de U et ne contenant donc aucun point de l'ensemble S tandis que le point ξ en tant que borne inférieure de l'ensemble S contient des points de S dans tout son voisinage. En particulier, $\xi \neq x$. Il est aussi évident que tout l'intervalle (x, ξ) est inclus dans U .

Faisons une construction analogue à gauche du point x ; nous aboutirons à un intervalle (η, x) contenu dans U dont l'extrémité gauche n'appartient pas à U (ou bien cet intervalle est toute la demi-droite $(-\infty, x)$).

Donc, nous avons construit, à partir d'un point donné $x \in U$, un intervalle (η, ξ) appartenant à l'ensemble U , dont chaque extrémité qui n'est pas à l'infini n'appartient pas à U . De tels intervalles sont appelés *intervalles composants de l'ensemble ouvert* U .

Si deux intervalles composants (η_1, ξ_1) et (η_2, ξ_2) ont un point commun x_0 , ils se confondent; en effet, par exemple, l'inégalité $\xi_1 < \xi_2$ est impossible puisque le point ξ_1 qui doit appartenir à l'ensemble U en tant que point intérieur de l'intervalle (x_0, ξ_2) ne le peut pas en tant qu'extrémité de l'intervalle (x_0, ξ_1) . Pour cette raison, tout l'ensemble U est la réunion de ses intervalles composants disjoints deux à deux. Une telle réunion ne peut être plus que dénombrable, car, dans tout intervalle composant, on peut choisir un point rationnel (1.75), or l'ensemble de tous les points rationnels est dénombrable (2.34). Notre proposition est définitivement démontrée.

§ 3.3. Suites convergentes et homéomorphisme

3.31. Suites convergentes. Soit une suite $\{x_n\} = \{x_1, \dots, x_n, \dots\}$ de points d'un espace métrique M . Nous dirons que cette suite *converge vers un point* $x \in M$ (et notons-le $x_n \rightarrow x$) lorsque, pour tout $\epsilon > 0$, on peut indiquer un N naturel tel que, pour tout $n \geq N$, on a l'inégalité

$$\rho(x, x_n) < \epsilon.$$

Autrement dit, la suite x_n converge vers x si tous les points de cette suite, à partir d'un point quelconque, appartiennent à toute boule centrée en x (il ne reste donc à l'extérieur de la boule qu'un nombre fini des points de la suite).

Le point x est alors appelé *limite* de la suite x_n , ce qu'on désigne par le symbole

$$x = \lim_{n \rightarrow \infty} x_n.$$

Le symbole $x_n \rightarrow x$ se lit « x_n converge vers x » ou, d'une façon plus détaillée, « la suite x_n converge vers x pour $n \rightarrow \infty$ sur la métrique ρ ». Si, pour une suite donnée x_n , il n'existe aucun point x pour lequel la relation $x_n \rightarrow x$ soit valable, la suite x_n est dite *divergente*.

3.32. Exemples.

a. Si M est la droite numérique, $M = R$, avec la métrique 3.13.a,

$$\rho(x, y) = |x - y|,$$

alors la définition 3.31 prend la forme suivante : une suite de nombres réels $x_1, x_2, \dots$ converge vers un nombre x si, pour tout $\varepsilon > 0$, on peut indiquer un N tel que, pour tout $n > N$, on a l'inégalité

$$|x - x_n| < \varepsilon.$$

Ainsi, la suite des points $x_n = \frac{1}{n}$ ($n = 1, 2, \dots$) sur la droite (avec sa métrique 3.13.a) converge vers le point $x = 0$; en effet, choisissons, pour un $\varepsilon > 0$ donné, un nombre naturel $N > \frac{1}{\varepsilon}$; pour tout $n \geq N$, nous avons $0 < x_n = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$, de sorte que les points respectifs x_n appartiennent à la boule de rayon ε et de centre $x = 0$.

b. Montrons que les relations $x = \lim_{n \rightarrow \infty} x_n, y = \lim_{n \rightarrow \infty} y_n$, dans un espace métrique M , impliquent la relation $\rho(x, y) = \lim_{n \rightarrow \infty} \rho(x_n, y_n)$ sur la droite numérique (cette propriété est parfois appelée « continuité de la distance »).

Pour un $\varepsilon > 0$ donné, choisissons un numéro N tel que, pour $n > N$, on ait les inégalités

$$\rho(x, x_n) < \frac{\varepsilon}{2}, \quad \rho(y, y_n) < \frac{\varepsilon}{2};$$

alors, d'après l'inégalité de quadrilatère 3.11 (1), on a pour les mêmes $n > N$:

$$|\rho(x, y) - \rho(x_n, y_n)| \leq \rho(x, x_n) + \rho(y, y_n) < \varepsilon,$$

ce qu'il fallait démontrer.

c. Bien que la théorie générale des limites soit exposée dans le chapitre 4, nous avons d'ores et déjà besoin de certaines propriétés simples des suites numériques convergentes. Montrons avant tout que, pour deux suites numériques x_n, y_n , les relations

$$x = \lim_{n \rightarrow \infty} x_n, \quad y = \lim_{n \rightarrow \infty} y_n, \quad x_n \leq y_n \quad (n = 1, 2, \dots)$$

impliquent l'inégalité

$$x \leq y.$$

En effet, supposons que $x > y$, et soit $\varepsilon = x - y > 0$. Trouvons un numéro N tel que, pour $n > N$, on ait les inégalités

$$|x_n - x| < \frac{\varepsilon}{2}, \quad |y_n - y| < \frac{\varepsilon}{2}.$$

Alors

$$x_n > x - \frac{\varepsilon}{2} = (y + \varepsilon) - \frac{\varepsilon}{2} > \left(y_n - \frac{\varepsilon}{2}\right) + \varepsilon - \frac{\varepsilon}{2} = y_n$$

en contradiction avec la supposition faite. La proposition est donc démontrée.

d. Prouvons que, pour deux suites numériques x_n, y_n , les relations

$$x = \lim_{n \rightarrow \infty} x_n, \quad y = \lim_{n \rightarrow \infty} y_n$$

impliquent

$$x + y = \lim_{n \rightarrow \infty} (x_n + y_n).$$

Trouvons, pour un $\varepsilon > 0$ donné, un numéro N de façon que, pour $n \geq N$, on ait les inégalités

$$|x - x_n| < \frac{\varepsilon}{2}, \quad |y - y_n| < \frac{\varepsilon}{2}.$$

Alors, pour les mêmes $n > N$, on a l'inégalité

$$\begin{aligned} |(x + y) - (x_n + y_n)| &= |(x - x_n) + (y - y_n)| \leq \\ &\leq |x - x_n| + |y - y_n| < \varepsilon, \end{aligned}$$

ce qu'il nous fallait.

e. Pour trois suites numériques x_n, y_n, z_n , les relations

$$x = \lim_{n \rightarrow \infty} x_n, \quad y = \lim_{n \rightarrow \infty} y_n, \quad z = \lim_{n \rightarrow \infty} z_n,$$

$$x_n + y_n \leq z_n \quad (n = 1, 2, \dots)$$

ont pour conséquence l'inégalité

$$x + y \leq z;$$

ce fait résulte immédiatement de c et d.

f. Soit M un espace réel m -dimensionnel muni de la métrique 3.14 (1): si $x(\xi_1, \dots, \xi_m)$, $y(\eta_1, \dots, \eta_m)$, alors

$$\rho(x, y) = \sqrt{\sum_{k=1}^m (\xi_k - \eta_k)^2}.$$

Démontrons que la convergence d'une suite de vecteurs $x_n = (\xi_{n1}, \dots, \xi_{nm})$ vers un vecteur $x = (\xi_1, \dots, \xi_m)$ équivaut à la convergence de m suites numériques:

$$\xi_{n1} \rightarrow \xi_1, \dots, \xi_{nm} \rightarrow \xi_m. \quad (1)$$

En effet, supposons remplies les relations (1). Pour un $\varepsilon > 0$ donné, on peut trouver un numéro N tel que, pour tout $n > N$, toutes les inégalités

$$|\xi_{n1} - \xi_1| < \frac{\varepsilon}{\sqrt{m}}, \dots, |\xi_{nm} - \xi_m| < \frac{\varepsilon}{\sqrt{m}}$$

soient remplies, de sorte que

$$\max_k |\xi_{nk} - \xi_k| < \frac{\varepsilon}{\sqrt{m}}.$$

Pour les mêmes $n > N$, en vertu de la deuxième inégalité 2.68 (7), nous avons

$$\rho(x_n, x) = \sqrt{\sum_{h=1}^m (\xi_{nh} - \xi_h)^2} \leq \sqrt{m} \max_h |\xi_{nh} - \xi_h| < \varepsilon,$$

de sorte que $x_n \rightarrow x$ sur la métrique de l'espace R_m .

Réciproquement, soit $x_n \rightarrow x$ sur la métrique de l'espace R_m . Pour un $\varepsilon > 0$ donné, trouvons un numéro N tel que, pour tout $n > N$, on ait l'inégalité

$$\rho(x_n, x) = \sqrt{\sum_{h=1}^m (\xi_{nh} - \xi_h)^2} < \varepsilon.$$

Dans ce cas, en vertu de la première inégalité 2.68 (7), on a :

$$\max_h |\xi_{nh} - \xi_h| \leq \rho(x_n, x) < \varepsilon,$$

de sorte que toute suite ξ_{nh} converge vers le nombre ξ_h ($h = 1, \dots, m$), ce que nous cherchions.

3.33.a. Vérifions que si une suite x_n est convergente, elle ne converge que vers un seul élément de l'espace M . Soit $x_n \rightarrow x$ et de même $x_n \rightarrow y$. Alors trouvons, pour un $\varepsilon > 0$ donné, un numéro N à partir duquel les deux inégalités

$$\rho(x, x_n) < \varepsilon, \quad \rho(y, x_n) < \varepsilon,$$

sont satisfaites, d'où, selon l'inégalité de triangle,

$$\rho(x, y) \leq \rho(x, x_n) + \rho(x_n, y) < 2\varepsilon.$$

Puisque $\varepsilon > 0$ est arbitrairement petit, on a $\rho(x, y) = 0$ d'après 1.57, d'où $x = y$ en vertu de l'axiome 3.11.b, ce qu'il fallait démontrer.

b. Montrons que si une suite $x_1, x_2, \dots$ est convergente, alors elle est bornée dans l'espace M ; autrement dit, les nombres $\rho(x_n, b)$, b étant un élément quelconque de l'espace M , forment sur la droite numérique un ensemble borné (cf. 3.12.a).

En effet, soit $p \equiv \lim_{n \rightarrow \infty} x_n$; pour un $\varepsilon > 0$, par exemple pour $\varepsilon = 1$, trouvons un nombre N tel que, pour $n > N$, on ait l'inégalité $\rho(x_n, p) < \varepsilon = 1$. Soit ensuite $D = \max\{\rho(x_1, p), \dots, \rho(x_N, p)\}$. Alors nous avons $\rho(x_n, p) \leq \max\{1, D\}$ pour tout $n = 1, 2, \dots$, ce qui démontre que la suite $x_1, x_2, \dots$ est bornée.

3.34. **H o m é o m o r p h i s m e d ' e s p a c e s m é t r i q u e s.** Dans bien des problèmes de l'analyse, c'est non pas la forme explicite de la métrique d'un espace qui importe mais seulement un moyen de reconnaître si telle ou telle suite est convergente ou non sur la métrique donnée. Ceci dit, introduisons la définition suivante :

a. Deux espaces métriques M' et M'' sont dits *homéomorphes* si l'on peut établir une correspondance biunivoque entre leurs points d'une telle façon qu'une convergence $x_n' \rightarrow x'$ dans M' implique la convergence des points correspondants $x_n'' \rightarrow x''$ dans M'' , et réciproquement, une convergence $x_n'' \rightarrow x''$ dans M'' implique la convergence des points correspondants $x_n' \rightarrow x'$ dans M' .

Une correspondance biunivoque décrite s'appelle *homéomorphisme*.

b. Le théorème suivant fournit un critère pour qu'une correspondance biunivoque entre les points de deux espaces métriques soit un homéomorphisme.

Théorème. Soit $\sim$ une application bijective d'un espace métrique M' avec ρ' (x', y') pour métrique sur un espace métrique M'' avec ρ'' (x'', y'') pour métrique. L'application $\sim$ est un homéomorphisme si, et seulement si, pour tout $x' \in M'$ et pour tout $\varepsilon' > 0$, il existe un $\varepsilon'' > 0$ tel que l'application inverse par rapport à $\sim$ transforme la boule $\{y'' \in M'' : \rho''(x'', y'') < \varepsilon''\}$ dans la boule $\{y' \in M' : \rho'(x', y') < \varepsilon'\}$, et réciproquement, pour tout $\varepsilon'' > 0$, il existe un $\varepsilon' > 0$ tel que l'application $\sim$ transforme la boule $\{y' \in M' : \rho'(x', y') < \varepsilon'\}$ dans la boule

$$\{y'' \in M'' : \rho''(x'', y'') < \varepsilon''\}.$$

Démonstration. Supposons que l'application bijective $\sim$ soit un homéomorphisme de sorte que, pour toute suite $x_n' \rightarrow x'$, nous avons $x_n'' \rightarrow x''$ et réciproquement. Fixons le point $x' \in M'$ et le nombre $\varepsilon' > 0$. Si le nombre ε'' mentionné dans l'énoncé du théorème n'existe pas, alors, pour tout $n = 1, 2, \dots$, la boule $\{y'' \in M'' : \rho''(x'', y'') < 1/n\}$ est transformée par l'application inverse par rapport à $\sim$ en l'ensemble contenant des points à l'extérieur de la boule $\{y' \in M' : \rho'(x', y') < \varepsilon'\}$. En particulier, pour tout $n = 1, 2, \dots$, on peut trouver un point $y_n'' \in M''$ pour lequel $\rho''(x'', y_n'') < \frac{1}{n}$, $\rho'(x', y_n'') \geq \varepsilon'$. Ainsi, $y_n'' \rightarrow x''$, or y_n'' ne converge pas vers x' , ce qui contredit la supposition. Donc, en partant d'un $\varepsilon' > 0$ donné on peut trouver ε'' mentionné dans l'énoncé si l'application $\sim$ est un homéomorphisme. En permutant M'' et M' dans le raisonnement précédent on peut trouver ε' d'après un $\varepsilon'' > 0$ donné.

Réciproquement, supposons que l'on puisse trouver un ε'' pour tout $x' \in M'$ et tout $\varepsilon' > 0$. Montrons que $x_n' \rightarrow x'$ implique $x_n'' \rightarrow x''$. Choisissons un $\varepsilon' > 0$ quelconque, trouvons un ε'' correspondant et un numéro N tel que, pour $n \geq N$, l'inégalité $\rho''(x_n'', x'') < \varepsilon''$ soit satisfaite. En appliquant $\sim$ et compte tenu de la supposition, nous obtenons $\rho'(x_n', x') < \varepsilon'$. Ainsi, $x_n' \rightarrow x'$. En permutant M'' et M' on obtient que $x_n'' \rightarrow x''$ implique $x_n' \rightarrow x'$. Le théorème est démontré.

c. On peut doter un ensemble E de métriques différentes en le transformant en de différents espaces métriques. Deux métriques

$\rho'(x, y)$ et $\rho''(x, y)$ sur un même espace E sont dites *homéomorphes* si l'application identique $x \sim x$ est un homéomorphisme des espaces métriques respectifs M' et M'' .

d. En appliquant le critère b à la situation décrite dans c on obtient le théorème suivant :

T h é o r è m e. Deux métriques ρ' et ρ'' données sur un même ensemble E sont homéomorphes si, et seulement si, pour tout $x \in E$ et tout $\varepsilon' > 0$ il existe un ε'' tel que la boule $\{y: \rho''(x, y) < \varepsilon''\}$ est incluse dans la boule $\{y: \rho'(x, y) < \varepsilon'\}$, et réciproquement, pour tout $\varepsilon'' > 0$, il existe un ε' tel que la boule $\{y: \rho'(x, y) < \varepsilon'\}$ est incluse dans la boule $\{y: \rho''(x, y) < \varepsilon''\}$.

e. Considérons à titre d'exemple trois métriques sur le produit cartésien M des espaces métriques M_1 et M_2 (3.16) :

$$\rho^{(1)}(x, y) = \max\{\rho_1(x_1, y_1), \rho_2(x_2, y_2)\};$$

$$\rho^{(2)}(x, y) = \rho_1(x_1, y_1) + \rho_2(x_2, y_2);$$

$$\rho^{(3)}(x, y) = \sqrt{\rho_1^2(x_1, y_1) + \rho_2^2(x_2, y_2)}.$$

Elles sont toutes homéomorphes sur M . Ceci découle des inégalités

$$\max\{a, b\} \leq \sqrt{a^2 + b^2} \leq a + b \leq 2 \max\{a, b\} \quad (1)$$

valables quels que soient les nombres non négatifs a et b . Plus en détail : soit, par exemple, $a \leq b$; alors

$$\begin{aligned} \max\{a, b\} &= b = \sqrt{b^2} \leq \sqrt{a^2 + b^2} \leq \sqrt{a^2 + 2ab + b^2} = \\ &= a + b \leq 2b = 2 \max\{a, b\}, \end{aligned}$$

d'où (1).

3.35. Nous introduirons plus bas une métrique sur l'axe numérique achevé $\bar{R}$ (1.91), cette métrique étant homéomorphe sur l'ensemble $-\infty < x < \infty$ à la métrique ordinaire 3.13.a. A cette fin, considérons d'abord une fonction spéciale.

a. Soit la fonction de la variable x , $-\infty < x < \infty$:

$$f(x) = \frac{x}{1+|x|}. \quad (1)$$

Complétons sa définition par les conditions

$$f(-\infty) = -1, \quad f(+\infty) = +1,$$

de sorte que la fonction $f(x)$ devienne définie sur la droite achevée $\bar{R}$. Le graphique de la fonction $f(x)$ est représenté fig. 3.1. Dans b, c et d, nous déduirons quelques inégalités concernant cette fonction dont nous aurons besoin.

b. Il est évident que $f(x) > 0$ pour $x > 0$ et $f(x) < 0$ pour $x < 0$, tandis que $f(0) = 0$. Ensuite, pour tout x fini, nous avons

$f(-x) = -f(x)$ et

$$|f(x)| = \frac{|x|}{1+|x|} < \frac{|x|+1}{1+|x|} = 1. \quad (2)$$

c. Montrons que, quels que soient x et y de R :

$$|f(x) - f(y)| \leq |x - y|. \quad (3)$$

Pour $x \geq 0, y \geq 0$, nous avons

$$|f(x) - f(y)| = \left| \frac{x}{1+x} - \frac{y}{1+y} \right| = \frac{|x-y|}{(1+x)(1+y)} \leq |x-y|.$$

De même, pour $x \leq 0, y \leq 0$:

$$|f(x) - f(y)| = \left| \frac{x}{1-x} - \frac{y}{1-y} \right| = \frac{|x-y|}{(1-x)(1-y)} \leq |x-y|.$$

Pour $x \leq 0, y \geq 0$, on a:

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x)| + |f(y)| = \\ &= \frac{|x|}{1+|x|} + \frac{y}{1+y} \leq |x| + y = -x + y = |x - y|; \end{aligned}$$

enfin, pour $x \geq 0, y \leq 0$, d'après ce qui précède:

$$|f(x) - f(y)| = |f(y) - f(x)| \leq |y - x| = |x - y|,$$

ce qu'il fallait démontrer.

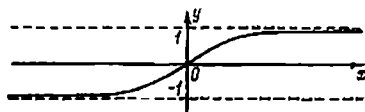


Fig. 3.1.

d. Supposons que, pour certains x et y et pour $0 < \delta < 1$, on ait

$$|f(x)| < 1 - \delta, \quad |f(y)| < 1 - \delta.$$

Alors

$$|x - y| \leq \frac{1}{\delta^2} |f(x) - f(y)|. \quad (4)$$

Pour démontrer, résolvons d'abord (1) par rapport à x . Pour $x \geq 0$ lorsque $|x| = x$, nous avons

$$x = \frac{f(x)}{1 - f(x)};$$

pour $x \leq 0$, lorsque $|x| = -x$, nous avons

$$x = \frac{f(x)}{1 + f(x)}.$$

Ainsi, pour $x \geq 0, y \geq 0$

$$\begin{aligned} |x-y| &= \left| \frac{f(x)}{1-f(x)} - \frac{f(y)}{1-f(y)} \right| = \\ &= \frac{|f(x)-f(y)|}{|1-f(x)||1-f(y)|} \leq \frac{1}{\delta^2} |f(x)-f(y)|. \end{aligned}$$

Si $x \leq 0, y \leq 0$, alors $-x \geq 0, -y \geq 0$, et d'après ce qui précède :

$$\begin{aligned} |x-y| &= |-y-(-x)| \leq \frac{1}{\delta^2} |f(-y)-f(-x)| = \\ &= \frac{1}{\delta^2} |f(x)-f(y)|. \end{aligned}$$

Si $x \leq 0, y \geq 0$, alors

$$\begin{aligned} |x-y| &= |x|+y = \frac{|f(x)|}{1+f(x)} + \frac{f(y)}{1-f(y)} \leq \\ &\leq \frac{1}{\delta^2} [|f(x)|+f(y)] = \frac{1}{\delta^2} |f(x)-f(y)|; \end{aligned}$$

si $x \geq 0, y \leq 0$, alors d'après ce qui précède :

$$|x-y| = |y-x| < \frac{1}{\delta^2} |f(y)-f(x)| = \frac{1}{\delta^2} |f(x)-f(y)|,$$

ce qu'il nous fallait.

e. Considérons maintenant, sur la droite numérique achevée $\bar{R}$ (1.91), la fonction numérique $r(x, y)$ d'un couple de points x, y :

$$r(x, y) = |f(x) - f(y)|. \quad (5)$$

Ici $f(x)$ est la fonction introduite dans a.

T h é o r è m e. *La fonction $r(x, y)$ (5) est une métrique sur $\bar{R}$: sur l'ensemble R des points finis $-\infty < x < \infty$, elle est homéomorphe à la métrique usuelle*

$$\rho(x, y) = |x - y|. \quad (6)$$

L'interprétation géométrique de la métrique $r(x, y)$ est évidente d'après la fig. 3.1. Notamment, on prend pour nouvelle distance $r(x, y)$ entre les points x et y de l'axe horizontal la longueur du segment correspondant $[f(x), f(y)]$ de l'axe vertical.

D é m o n s t r a t i o n d u t h é o r è m e. Vérifions les axiomes 3.11.a-c. Il résulte directement de la définition (5) que l'axiome b (celui de symétrie) est satisfait. Pour l'axiome a ($r(x, y) = 0$ pour $x = y$ et positif pour $x \neq y$), cela découle aussi de la définition et de l'inégalité (3). L'axiome c (de triangle) est satisfait d'après

l'inégalité

$$r(x, z) = |f(x) - f(z)| \leq |f(x) - f(y)| + |f(y) - f(z)| = \\ = r(x, y) + r(y, z).$$

Prouvons que les métriques $r(x, y)$ et $\rho(x, y)$ sont homéomorphes sur l'ensemble R des points finis.

Soit $x_n \rightarrow x$ sur la métrique (6). Pour un $\varepsilon > 0$ donné, trouvons un numéro N tel que, pour $n \geq N$, l'inégalité $|x_n - x| < \varepsilon$ soit satisfaite. Alors, d'après l'inégalité (3) pour les mêmes $n \geq N$, on a

$$r(x_n, x) = |f(x_n) - f(x)| \leq |x_n - x| < \varepsilon,$$

de sorte que $x_n \rightarrow x$ sur la métrique (5) aussi.

Soit maintenant $x_n \rightarrow x$ sur la métrique (5). Nous avons $|f(x)| < 1$; posons $|f(x)| = 1 - 2\delta$, $0 < \delta < \frac{1}{2}$. Pour un $\varepsilon > 0$ donné, trouvons un numéro N de façon à avoir, pour $n \geq N$:

$$r(x_n, x) = |f(x_n) - f(x)| < \min(\varepsilon, \delta) \cdot \delta^2.$$

Alors, pour les mêmes $n \geq N$:

$$|f(x_n)| \leq |f(x)| + \min(\varepsilon, \delta) \cdot \delta^2 \leq |f(x)| + \delta = 1 - \delta;$$

donc, étant donné que $|f(x)| = 1 - 2\delta < 1 - \delta$, on peut appliquer l'inégalité (4):

$$\rho(x_n, x) = |x_n - x| \leq \frac{1}{\delta^2} |f(x_n) - f(x)| \leq \min(\varepsilon, \delta) \leq \varepsilon.$$

Donc $x_n \rightarrow x$ sur la métrique (6) aussi. Le théorème est démontré.

f. Décrivons dans l'espace $\bar{R}$ muni de la métrique (5) la boule d'un rayon $c > 0$ centrée en ∞ . Par définition, elle est l'ensemble des $x \in \bar{R}$ qui vérifient l'inégalité

$$|f(\infty) - f(x)| = \left| 1 - \frac{x}{1+x} \right| \leq c. \quad (7)$$

Bornons-nous au cas le plus important où c est petit, $c \leq 1$. Alors, comme les valeurs de $f(x)$ sont négatives pour $x < 0$ et que $f(\infty) - f(x) > 1$, l'inégalité (7) ne peut être vérifiée que pour les valeurs $x \geq 0$ et (7) prend la forme

$$1 - \frac{x}{1+x} = \frac{1}{1+x} \leq c,$$

ou

$$x \geq \frac{1}{c} - 1.$$

Définitivement, la boule en question $\{x: r(x, \infty) \leq c\}$ est l'intervalle $\frac{1}{c} - 1 \leq x \leq \infty$. D'une façon analogue, la boule de rayon

$c \leq 1$ centrée en $-\infty$ est l'intervalle $-\infty \leq x \leq -\frac{1}{c} + 1$.

En particulier, la suite des points 1, 2, 3, . . . dans l'espace $\bar{R}$ appartient, à partir d'un certain numéro, à toute boule de centre ∞ , de sorte que nous avons dans l'espace $\bar{R}$

$$\lim_{n \rightarrow \infty} n = \infty,$$

tandis que, dans l'espace R , ladite suite n'a, évidemment, aucune limite.

g. Il est à signaler que l'espace $\bar{R}$ doté de la métrique (5) est isométrique à l'intervalle $[-1, 1]$ avec sa métrique usuelle. La correspondance $\bar{R} \sim [-1, 1]$ se réalise suivant la règle $x \sim f(x)$; elle est biunivoque par construction, la distance entre deux points $x, y \in \bar{R}$ pour la métrique (5) étant égale à la distance usuelle entre les points correspondants $f(x)$ et $f(y)$.

Une fois les extrémités enlevées ($-\infty$ et ∞ pour l'espace $\bar{R}$ et $-1, +1$ pour l'intervalle fermé $[-1, 1]$), on obtient l'isométrie de l'axe réel R doté de la métrique (5) et de l'intervalle ouvert $(-1, 1)$ avec sa métrique usuelle. En appliquant e on voit que l'espace métrique R avec sa métrique usuelle est homéomorphe à l'intervalle ouvert $(-1, 1)$, le fait qui peut, d'ailleurs, être vérifié directement.

§ 3.4. Valeurs d'adhérence et points d'accumulation

3.41. Soit toujours une suite $\{x_n\}$ de points d'un espace métrique M . Nous dirons qu'un point $y \in M$ est *valeur d'adhérence* de la suite $\{x_n\}$ si, pour tout $\varepsilon > 0$ et tout N naturel, on peut trouver un numéro $n \geq N$ pour lequel

$$\rho(y, x_n) < \varepsilon.$$

Autrement dit, un point y est une valeur d'adhérence d'une suite x_n si toute boule centrée en y contient des points de la suite ayant les numéros aussi grands que l'on veut, mais pas nécessairement tous, contrairement au cas d'une suite convergente. Une suite convergente $x_n \rightarrow x$ a le point x et seulement lui pour une valeur d'adhérence (la démonstration est analogue à celle de 3.33.a). Une suite divergente peut ne pas avoir de valeurs d'adhérence ou bien en avoir une quantité arbitraire. Ainsi, la suite des points $x_n = (-1)^n \left(1 + \frac{1}{n}\right)$ sur la droite numérique (munie de sa métrique usuelle) a deux valeurs d'adhérence -1 et $+1$ (et ne converge vers aucune d'elles); la suite $x_n = n^{(-1)^n}$ possède une seule valeur d'adhérence $y = 0$ et ne conver-

ge pas vers elle; et si l'on range tous les points rationnels sur la droite en une suite (2.34), alors tout point de la droite sera une valeur d'adhérence de cette suite.

3.42. Si une suite $x_1, \dots, x_n, \dots$ a une valeur d'adhérence y , on peut extraire de la suite x_n une sous-suite convergente vers y . A savoir, si y est une valeur d'adhérence de la suite x_n , alors, pour tout m , il existe dans la suite x_n un point x_{n_m} ($n_m > n_{m-1}$) pour lequel $\rho(y, x_{n_m}) < \frac{1}{m}$; la sous-suite composée des points $x_{n_1}, x_{n_2}, \dots$ converge, évidemment, vers le point y . Réciproquement, si une suite x_n admet une sous-suite x_{n_m} ($m = 1, 2, \dots$) convergente vers un point $y \in M$, alors ce point y est bien une valeur d'adhérence de toute la suite x_n . Nous aboutissons ainsi à la deuxième définition d'une valeur d'adhérence: *un point y est une valeur d'adhérence d'une suite x_n si la suite x_n possède une sous-suite qui converge vers le point y .*

3.43. La deuxième définition d'une valeur d'adhérence n'est basée que sur la notion de suite convergente. Vu qu'un homéomorphisme d'un espace métrique M dans un espace métrique M' conserve la convergence d'une suite, nous arrivons à la conclusion suivante: *si un point $y \in M$ est une valeur d'adhérence d'une suite $x_n \in M$ et un espace M' est homéomorphe à l'espace M , alors le point $y' \in M'$ qui correspond au point $y \in M$ par l'homéomorphisme en question est une valeur d'adhérence pour la suite des points correspondants $x'_n \in M'$.*

En particulier, lorsqu'on introduit, sur un même ensemble M , deux métriques différentes mais homéomorphes ρ et ρ' et qu'un point $y \in M$ est une valeur d'adhérence d'une suite x_n sur la métrique ρ , alors il est une valeur d'adhérence de la suite x_n aussi bien sur la métrique ρ' .

3.44.a. Soit A un sous-ensemble d'un espace métrique M . On dit qu'un point $y \in M$ est un *point d'accumulation de l'ensemble A* si tout voisinage $V_r(y) = \{x: \rho(x, y) < r\}$ du point y contient un point $x \in A$ autre que y .

Les propositions 3.42 et 3.43 concernant les valeurs d'adhérence restent valables pour les points d'accumulation. Si y est un point d'accumulation d'un ensemble $A \subset M$, alors on peut former (de même qu'en 3.42) une suite de points distincts de l'ensemble A qui converge vers y ; réciproquement, si l'on peut choisir une suite de points distincts de l'ensemble A qui converge vers un point $y \in M$, alors ce point y est un point d'accumulation pour l'ensemble A . Il en résulte la conservation des points d'accumulation de l'ensemble A par tout homéomorphisme de l'espace métrique M (en vertu de 3.43).

b. E x e m p l e. Soit A un ensemble majoré (minoré) sur l'axe réel et supposons que $\xi = \sup A$ ($\eta = \inf A$) n'appartienne pas à

l'ensemble A ; montrons que $\xi(\eta)$ est un point d'accumulation de l'ensemble A .

Faisons la démonstration pour le point $\xi = \sup A$. Selon la définition même de la borne supérieure, pour tout $n = 1, 2, \dots$, il existe un point $x_n \in A$ tel que $\xi - \frac{1}{n} < x \leq \xi$. Puisque $\xi \notin A$ par hypothèse, on a $x_n \neq \xi$. Le nombre ξ est donc un point d'accumulation de l'ensemble A , ce qu'il fallait démontrer.

3.45. Il serait bien d'avoir des propositions générales permettant de juger de l'existence de points d'accumulation pour une classe assez vaste d'ensembles. Le théorème suivant est une proposition de ce genre :

Théorème de Bolzano-Weierstrass. *Tout ensemble infini de points du segment $[a, b] \subset R$ a au moins un point d'accumulation.*

Démonstration. Si un ensemble infini E fait partie d'un intervalle $[c, d]$, alors l'une au moins des moitiés $\left[c, \frac{c+d}{2}\right]$ et $\left[\frac{c+d}{2}, d\right]$ de l'intervalle $[c, d]$ contient un sous-ensemble infini de l'ensemble E . Cette considération élémentaire nous permet de construire, à partir de l'intervalle $[a, b] = \Delta_1$, sur lequel un ensemble infini de points A est donné, une suite des intervalles fermés emboîtés $\Delta_1 \supset \Delta_2 \supset \dots$, où chaque intervalle suivant est la moitié du précédent et contient un sous-ensemble infini de l'ensemble A . D'après le principe de Cantor 1.81, ces intervalles ont un point commun x_0 ; montrons qu'il est un point d'accumulation de l'ensemble A . Choisissons un intervalle V de longueur $\delta > 0$ et soit x_0 son milieu. Soit un n tel que la longueur de l'intervalle Δ_n est inférieure à $\frac{\delta}{2}$; cet intervalle est inclus dans V tout en contenant le point x_0 . Avec l'intervalle Δ_n , l'intervalle V enferme une infinité de points de l'ensemble A ; donc x_0 est un point d'accumulation de l'ensemble A , ce qu'il fallait démontrer.

3.46. Une proposition analogue est valable pour les suites :

Théorème (théorème de Bolzano-Weierstrass pour les suites). *Toute suite de points d'un intervalle $[a, b]$ possède au moins une valeur d'adhérence.*

La démonstration répète celle de 3.45, compte tenu du fait que les points d'une suite peuvent se répéter.

3.47. Sur toute la droite numérique R , il y a des ensembles infinis sans points d'accumulation (par exemple 1, 2, 3, ...). Mais sur la droite achevée $\bar{R}$ métrisée suivant la règle 3.35.e, tout

ensemble infini possède toujours un point d'accumulation; ceci résulte du fait que $\bar{R}$ est isométrique à l'intervalle $[-1, 1]$ avec sa métrique usuelle (3.35.g) et que la propriété d'un point d'être point d'accumulation est conservée par une application isométrique (et même par un homéomorphisme, ce qu'on a vu dans 3.43).

3.48. La proposition suivante fournit une condition suffisante d'absence de valeurs d'adhérence:

Une suite $x_1, x_2, \dots$ pour laquelle toutes les distances $\rho(x_m, x_n)$ sont minorées par une constante positive C , de sorte que $\rho(x_m, x_n) \geq C$, n'a aucune valeur d'adhérence.

En effet, soit y une valeur d'adhérence de la suite x_n ; il existe donc, dans cette suite, des points x_m et x_n , $m \neq n$, tels que $\rho(y, x_m) < \frac{C}{2}$, $\rho(y, x_n) < \frac{C}{2}$, d'où

$$\rho(x_m, x_n) \leq \rho(x_m, y) + \rho(y, x_n) < C,$$

en contradiction avec la supposition.

§ 3.5. Ensembles fermés

3.51. Définition. Un ensemble $F \subset M$ est dit *fermé* s'il contient tous ses points d'accumulation.

Exemples. a. Un intervalle $a \leq x \leq b$ est fermé sur la droite numérique et l'intervalle $a \leq x < b$ n'est pas fermé, son point d'accumulation b ne lui appartenant pas.

b. Dans tout espace métrique, chaque boule

$$U = \{x: \rho(x, x_0) \leq r\}$$

est un ensemble fermé (c'est pour cette raison qu'elle est appelée *boule fermée*).

En effet, soit x_1 un point quelconque n'appartenant pas à la boule U , de sorte que $\rho(x_0, x_1) = r_1 > r$. La boule de centre x_1 et de rayon $\frac{1}{2}(r_1 - r)$ ne contient aucun point de la boule U : s'il y en avait un, alors, après l'avoir désigné par z , on aurait

$$\rho(x_0, x_1) \leq \rho(x_0, z) + \rho(x_1, z) \leq r + \frac{1}{2}(r_1 - r) = \frac{1}{2}(r_1 + r) < r_1,$$

en contradiction avec l'égalité $\rho(x_0, x_1) = r_1$. Donc, le point x_1 ne peut être un point d'accumulation de l'ensemble U .

3.52. Les ensembles fermés dans un espace métrique M sont liés d'une façon étroite aux ensembles ouverts de cet espace. Notamment, le théorème suivant a lieu:

Théorème. Dans un espace métrique M , le complémentaire U d'un ensemble fermé F est toujours ouvert. Le complémentaire F d'un ensemble ouvert U est toujours fermé.

Démonstration. Soient F un ensemble fermé et U son complémentaire; montrons que U est ouvert. Considérons son point quelconque $x_0 \in U$; il faut montrer qu'il existe une boule définie par l'inégalité

$$\rho(x, x_0) < r$$

et incluse dans l'ensemble U .

Raisonnons par l'absurde. Supposons que toute boule centrée en x_0 contienne des points de l'ensemble F . Alors, d'après la définition d'un point d'accumulation, le point x_0 est un point d'accumulation de l'ensemble F . Puisque F est fermé, nous avons $x_0 \in F$, ce qui contredit la supposition $x_0 \in U$. Donc U est ouvert.

Passons à la deuxième partie du théorème. Soit U un ensemble ouvert et F son complémentaire; montrons que F est fermé. Tout point x_0 appartenant à U y est contenu avec une boule et ne peut donc être un point d'accumulation de l'ensemble F . Ainsi, les points d'accumulation de l'ensemble F ne peuvent être contenus que dans F lui-même, donc F est fermé. Le théorème est démontré.

3.53.a. En nous rappelant 3.23 nous obtenons une description générale de tous les ensembles fermés sur la droite $-\infty < x < \infty$: tout ensemble fermé de la droite s'obtient en enlevant de la droite une famille finie ou dénombrable d'intervalles ouverts disjoints deux à deux. Les intervalles enlevés qui composent le complémentaire ouvert sont appelés *intervalles contigus* à l'ensemble fermé donné.

b. Un ensemble fermé A qui est majoré (minoré) contient toujours sa borne supérieure (inférieure); sinon cette borne serait un point d'accumulation de A (3.44.b), donc l'ensemble se trouverait non fermé.

3.54. Vu les propriétés connues des ensembles ouverts dans un espace métrique (3.22) et la relation entre les ensembles ouverts et fermés que l'on vient de discuter, nous pouvons affirmer que la réunion d'un nombre fini d'ensembles fermés ainsi que l'intersection d'une famille quelconque d'ensembles fermés sont des ensembles fermés.

En effet, soient F_ν des ensembles fermés, ν parcourant un ensemble d'indices, et U_ν les ensembles ouverts complémentaires correspondants. D'après la formule 2.12 (1), nous avons

$$C\left[\bigcap_\nu F_\nu\right] = \sum_\nu C F_\nu = \sum_\nu U_\nu.$$

L'ensemble $\sum_\nu U_\nu$ est ouvert, donc son complémentaire $\left[\bigcap_\nu F_\nu\right]$ est fermé.

Si v parcourt un ensemble fini d'indices, alors $\prod_v U_v$ est ouvert d'après 3.22; étant donné 2.12 (2), nous avons

$$C \sum_v F_v = \prod_v C F_v = \prod_v U_v;$$

par conséquent, $\sum F_v$ est fermé, ce qu'il fallait démontrer.

§ 3.6. Ensembles partout denses. Fermetures

3.61. Ensembles partout denses. On dit qu'un ensemble A dans un espace métrique M est *partout dense par rapport à un ensemble* $B \subset M$ si tout point $x \in B$ ou bien appartient lui-même à A , ou bien en est un point d'accumulation. En d'autres termes, A est partout dense par rapport à B si toute boule centrée en un point $x \in B$ contient un point $y \in A$.

Si de plus $A \subset B$, on dit que A est *partout dense dans* B . Ainsi, l'ensemble des points rationnels est partout dense sur la droite numérique $-\infty < x < \infty$, l'ensemble des points $(r_1, \dots, r_n)$ de coordonnées rationnelles est partout dense dans l'espace euclidien n -dimensionnel.

3.62. La propriété d'être « partout dense » est transitive au sens suivant: si un ensemble A est partout dense par rapport à un ensemble B , et l'ensemble B est, à son tour, partout dense par rapport à un ensemble C , alors A est partout dense par rapport à C aussi. En effet, pour un $\varepsilon > 0$ donné et pour un point donné $z \in C$, on peut trouver un point $y \in B$ tel que $\rho(y, z) < \frac{\varepsilon}{2}$; ensuite, on peut trouver un point $x \in A$ tel que $\rho(x, y) < \frac{\varepsilon}{2}$; définitivement, pour tout $\varepsilon > 0$, on peut trouver en partant d'un point donné $z \in C$ un point $x \in A$ tel que $\rho(x, z) \leq \rho(x, y) + \rho(y, z) < \varepsilon$. Ceci veut dire que A est partout dense par rapport à C .

3.63. Fermeture. Soit A un ensemble dans un espace métrique M . L'ensemble des points de A et de tous ses points d'accumulation s'appelle *fermeture* de l'ensemble A et se note $\bar{A}$. Si A est un ensemble fermé, c'est-à-dire que A contient tous ses points d'accumulation, alors $\bar{A} = A$. En général, $\bar{A} \supset A$. Si $\bar{A} = A$, cela veut dire que tous les points d'accumulation de l'ensemble A appartiennent à A et que, par conséquent, A est fermé. Donc, les affirmations « A est fermé » et « $\bar{A} = A$ » sont équivalentes. Si $\bar{A} \subset A$, alors, puisque on a toujours $\bar{A} \supset A$, il vient $\bar{A} = A$, de sorte que A est fermé.

3.64. Montrons que la *fermeture de tout ensemble A est un ensemble fermé*. Remarquons d'abord que, par définition, tout ensemble est partout dense dans sa fermeture. En particulier, A est partout dense dans $\bar{A}$, $\bar{A}$ est partout dense dans sa fermeture $\bar{\bar{A}}$. Il en résulte, vu la transitivité de la propriété, que A est partout dense dans $\bar{\bar{A}}$, c'est-à-dire que tout point de l'ensemble $\bar{\bar{A}}$ est ou bien un point de l'ensemble A ou bien un de ses points d'accumulation. Par suite, $\bar{\bar{A}} \subset \bar{A}$, donc $\bar{A}$ est fermé.

Ainsi, la fermeture de l'ensemble des points rationnels de l'axe $-\infty < x < \infty$ se confond avec tout l'axe.

3.65. Montrons que la *fermeture $\bar{A}$ d'un ensemble borné A est un ensemble borné et que $\text{diam } \bar{A} = \text{diam } A$* .

Il suit de $A \subset \bar{A}$ que $\text{diam } A \leq \text{diam } \bar{A}$. Ensuite, étant donnés deux points quelconques $\bar{x}$ et $\bar{y}$ de l'ensemble $\bar{A}$ et un $\varepsilon > 0$, choisissons des points $x \in A$, $y \in A$ tels que $\rho(x, \bar{x}) < \frac{\varepsilon}{2}$, $\rho(y, \bar{y}) < \frac{\varepsilon}{2}$. Alors

$$\rho(\bar{x}, \bar{y}) \leq \rho(\bar{x}, x) + \rho(x, y) + \rho(y, \bar{y}) < \frac{\varepsilon}{2} + \text{diam } A + \frac{\varepsilon}{2},$$

d'où

$$\text{diam } \bar{A} = \sup \rho(\bar{x}, \bar{y}) \leq \varepsilon + \text{diam } A$$

et, $\varepsilon > 0$ étant arbitraire :

$$\text{diam } \bar{A} \leq \text{diam } A.$$

Compte tenu de ce qui précède, nous avons bien $\text{diam } \bar{A} = \text{diam } A$.

3.66. Soient F un ensemble fermé et $G \supset F$ un domaine dans un espace métrique M . Montrons qu'il existe un domaine $H \supset F$ contenu dans le domaine G avec sa fermeture et tel que

$$F \subset H \subset \bar{H} \subset G.$$

A cette fin, désignons par $d(x)$ la distance d'un point $x \in G$ à l'extérieur du domaine G , c'est-à-dire

$$d(x) = \inf_{y \in M-G} \rho(x, y).$$

Pour tout $x \in G$, la quantité $d(x)$ est positive car le point x appartient au domaine G avec son voisinage. Définissons l'ensemble H comme réunion de toutes les boules ouvertes de centres $x \in F$ et de rayons $\frac{1}{2} d(x)$. Il est évident que $F \subset H$; d'après 3.22, H est un domaine. Montrons que la fermeture $\bar{H}$ de l'ensemble H est conte-

nue dans le domaine G . Supposons le contraire: il existe un point $y \in \overline{H} \cap (M - G)$. Trouvons une suite $z_n \in H$, $z_n \rightarrow y$ et une suite $x_n \in F$ avec $\rho(x_n, z_n) \leq \frac{1}{2} d(x_n)$. Nous avons:

$$d(x_n) \leq \rho(x_n, y) \leq \rho(x_n, z_n) + \rho(z_n, y) \leq \frac{1}{2} d(x_n) + \rho(z_n, y).$$

Il en résulte

$$d(x_n) \leq 2\rho(z_n, y), \quad \rho(x_n, y) \leq 2\rho(z_n, y),$$

de sorte que $y = \lim x_n \in F \subset G$, ce qui contredit la supposition $y \in M - G$. Donc, $\overline{H} \subset G$.

3.67. En particulier, si F est borné (3.12.a), alors la fonction $d(x)$ admet l'estimation (x_0 est un point fixe de F):

$$\begin{aligned} d(x) &= \inf_{y \in M-G} \rho(x, y) \leq \\ &\leq \inf_{y \in M-G} [\rho(x, x_0) + \rho(x_0, y)] \leq \text{diam } F + d(x_0). \end{aligned} \quad (1)$$

Par conséquent, l'ensemble H construit dans 3.66 est borné si F l'est, parce que la définition de H et (1) impliquent

$$\text{diam } H \leq \text{diam } F + 2 \sup_{x \in F} d(x) \leq 3 \text{diam } F + 2d(x_0).$$

§ 3.7. Espaces complets

3.71.a. Une suite $x_1, x_2, \dots$ de points d'un espace métrique M est appelée *suite de Cauchy* si, pour tout ε , il existe un numéro N tel que l'inégalité $\rho(x_v, x_\mu) \leq \varepsilon$ est satisfaite dès que $v \geq N$ et $\mu \geq N$.

b. Toute suite convergente $x_1, x_2, \dots \rightarrow x$ est une suite de Cauchy. En effet, d'après l'inégalité de triangle on a:

$$\rho(x_v, x_\mu) \leq \rho(x_v, x) + \rho(x, x_\mu),$$

et, lorsque $x_v \rightarrow x$, le second membre devient plus petit que tout $\varepsilon > 0$ donné, pour v et μ assez grands.

c. Toute suite de Cauchy $x_1, x_2, \dots$ est bornée. Pour le démontrer, choisissons un numéro N de façon à satisfaire l'inégalité $\rho(x_v, x_\mu) \leq 1$ avec $v \geq N$ et $\mu \geq N$. Nous avons alors, pour tout $v = 1, 2, \dots$:

$$\rho(x_v, x_N) \leq \begin{cases} \max\{\rho(x_1, x_N), \dots, \rho(x_{N-1}, x_N)\} & (v \leq N), \\ 1 & (v \geq N), \end{cases}$$

de sorte que les distances des points $x_1, x_2, \dots$ au point x_N sont bornées par une constante fixe.

d. **Définition.** Un espace métrique M est dit *complet* si toute suite de Cauchy de ses points converge vers un élément de cet espace.

3.72. Exemples.

a. Montrons que l'axe réel R doté de la métrique

$$\rho(x, y) = |x - y|$$

est un espace métrique complet.

Autrement dit, pour qu'une suite numérique $x_1, x_2, \dots$ soit convergente, il faut et il suffit que, pour tout $\varepsilon > 0$, il existe un numéro N tel que l'inégalité $|x_n - x_m| < \varepsilon$ soit satisfaite dès que $n > N$, $m > N$.

Ainsi énoncé, le théorème a est nommé « critère de Cauchy ». Pour le démontrer, nous procédons comme suit :

Soit $x_1, x_2, \dots$ une suite de Cauchy de nombres réels. Elle est bornée en vertu de 3.71.b ; pour tout $m = 1, 2, \dots$, posons

$$a_m = \inf_{n \geq m} x_n, \quad b_m = \sup_{n \geq m} x_n.$$

Il est clair que les inégalités $a_m \leq a_{m+1}$, $b_m \geq b_{m+1}$ sont valables pour tout m , d'où $[a_m, b_m] \supset [a_{m+1}, b_{m+1}]$.

D'après le principe de Cantor 1.81, il existe un point $p \in R$ commun à tous les intervalles $[a_m, b_m]$, $m = 1, 2, \dots$. Montrons que $p = \lim_{n \rightarrow \infty} x_n$. Pour un $\varepsilon > 0$ donné, trouvons un numéro N de façon à avoir, pour $n \geq N$, $m \geq N$:

$$\rho(x_n, x_m) = |x_n - x_m| \leq \varepsilon.$$

Fixons ici $m = N$ et posons successivement $n = N + 1, N + 2, \dots$. Toutes les valeurs de x_n étant différentes de x_N au plus de ε , pour $n > N$, il en est de même des nombres a_N et b_N . Le nombre p appartient à l'intervalle $[a_N, b_N]$, nous avons donc, pour tous les $n > N$:

$$|p - x_n| \leq b_N - a_N = (b_N - x_N) + (x_N - a_N) \leq 2\varepsilon.$$

Ainsi, $x_n \rightarrow p$, ce qu'il fallait démontrer.

b. L'intervalle $(0, 1)$ est un espace métrique pour la métrique usuelle de la droite numérique. La suite $\frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$ est une suite de Cauchy dans cet espace métrique, mais n'y possède pas de limite ; par conséquent, l'intervalle $(0, 1)$ n'est pas un espace complet.

Cet exemple montre, entre autres, que la propriété d'un espace d'être complet dépend essentiellement du choix de la métrique et peut ne pas être conservée en passant à une métrique homéomorphe ; en effet, les espaces métriques des exemples a et b sont homéomorphes (3.35.g), mais l'un d'eux est complet et l'autre non.

c. Montrons qu'un espace réel m -dimensionnel R_m muni de la métrique 3.14 (1) est complet. Soit $x_n = (\xi_{n1}, \dots, \xi_{nm})$ ($n = 1, 2, \dots$) une suite de Cauchy de vecteurs de l'espace R_m . Choisissons, pour un $\varepsilon > 0$ donné, un numéro N de façon à avoir, pour $p > N$, $q > N$:

$$|x_p - x_q| = \sqrt{\sum_{k=1}^m (\xi_{pk} - \xi_{qk})^2} < \varepsilon.$$

D'après l'inégalité 2.68 (7), nous avons pour tout k et pour $p, q > N$:

$$|\xi_{pk} - \xi_{qk}| \leq \sqrt{\sum_{k=1}^m (\xi_{pk} - \xi_{qk})^2} < \varepsilon.$$

Nous voyons que toute suite numérique ξ_{pk} ($k = 1, \dots, m$) est une suite de Cauchy de la droite numérique. En appliquant a, nous établissons l'existence des limites

$$\xi_k = \lim_{p \rightarrow \infty} \xi_{pk} \quad (k = 1, \dots, m).$$

Formons le vecteur $x = (\xi_1, \dots, \xi_m)$. Toujours d'après 2.68 (7):

$$|x - x_p| = \sqrt{\sum_{k=1}^m (\xi_k - \xi_{pk})^2} \leq \sqrt{m} \max_k |\xi_k - \xi_{pk}|. \quad (1)$$

A présent choisissons, pour un $\varepsilon > 0$ donné, un numéro N tel que, pour $p > N$, toutes les inégalités

$$|\xi_k - \xi_{pk}| < \frac{\varepsilon}{\sqrt{m}} \quad (k = 1, \dots, m)$$

soient satisfaites.

Alors, pour les mêmes valeurs de p , on a aussi

$$|x - x_p| < \varepsilon,$$

ce qui est évident de (1); il en découle que la suite des vecteurs x_n converge vers le vecteur x . Donc, l'espace R_m est bien complet.

3.73.a. Soit M un espace métrique complet qui est une partie (avec la même métrique) d'un espace métrique P . Montrons que M est fermé dans P . Soient $y \in P$ un point d'accumulation de l'ensemble M et x_n une suite de points de M qui converge vers le point y . Comme x_n est une suite de Cauchy (3.71.b) et que l'espace M est complet, il existe dans M un point $z = \lim_{n \rightarrow \infty} x_n$. En vertu de l'unicité d'une

limite (3.33.a), $y = z \in M$, ce qu'il nous fallait.

b. Réciproquement, si l'on considère un sous-ensemble fermé A d'un espace métrique complet M en tant qu'espace métrique (avec la métrique définie sur M), alors A est un espace métrique complet.

En effet, toute suite de Cauchy $y_n \in A$ converge dans M (M étant complet) et sa limite appartient à l'ensemble A , ce dernier étant fermé.

e. En particulier, tout intervalle $[a, b]$ de la droite numérique est un espace complet, vu qu'il est un sous-ensemble fermé (3.51.a) d'un espace métrique complet (3.72.a).

3.74. Le principe de Cantor d'intervalles emboîtés (1.81) est valable pour la droite numérique: tout système d'intervalles fermés emboîtés a un point commun. Pour un espace métrique complet, on peut signaler diverses propositions analogues.

a. Soit Q une famille de sous-ensembles non vides $A, B, \dots$ d'un ensemble M qui possèdent la propriété suivante: quels que soient deux sous-ensembles de la famille, l'un est inclus dans l'autre. Une telle famille Q est appelée *système de sous-ensembles emboîtés*.

L e m m e. *Soit Q un système de sous-ensembles emboîtés $A, B, \dots$ dans un espace métrique complet M . S'il y a parmi eux des sous-ensembles d'un diamètre (3.12.a) aussi petit que l'on veut, alors il existe un point $p \in M$ tel que, pour tout son voisinage $V_r(p) = \{x: p(x, p) < r\}$, on peut trouver un ensemble $A \in Q$ inclus dans $V_r(p)$.*

D é m o n s t r a t i o n. Par hypothèse, pour tout $m = 1, 2, 3, \dots$, il existe un ensemble $A = A_m \in Q$ tel que $\text{diam } A_m \leq \frac{1}{m}$. Soit x_m un point quelconque de A_m . Puisque, pour $n > m$, on a ou bien $A_n \subset A_m$, ou bien $A_m \subset A_n$, alors $p(x_n, x_m) \leq \frac{1}{m}$, de sorte que la suite $x_1, x_2, \dots$ est de Cauchy. Soit $p = \lim_{m \rightarrow \infty} x_m$; montrons que ce point p satisfait aux exigences du lemme. Soit un $\varepsilon > 0$ arbitraire; trouvons un n tel que $p(p, x_n) \leq \frac{\varepsilon}{2}$ et $\text{diam } A_n \leq \frac{\varepsilon}{2}$. Alors, pour tout $x \in A_n$, on a

$$p(p, x) \leq p(p, x_n) + p(x_n, x) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

de sorte que $A_n \subset V_\varepsilon(p)$, ce qu'il fallait démontrer.

b. **R e m a r q u e.** *Dans l'hypothèse du lemme a, il ne peut exister qu'un seul point p satisfaisant aux exigences du lemme.* En effet, s'il en existe un autre, q , de sorte que $p(p, q) = 2\varepsilon > 0$, alors les boules $V_\varepsilon(p)$ et $V_\varepsilon(q)$ sont disjointes; de plus, d'après le lemme a, il existe dans le système Q des ensembles A et B tels que $A \subset V_\varepsilon(p)$, $B \subset V_\varepsilon(q)$; or, ceci contredit le fait que l'un des ensembles A et B est inclus dans l'autre.

c. **L e m m e.** *Complétons l'hypothèse du lemme a par la supposition que tous les ensembles du système Q soient fermés. Alors le point p*

satisfaisant aux exigences du lemme appartient à tout ensemble du système Q .

Démonstration. Supposons que le point p n'appartienne pas à un ensemble B . B étant fermé, il existe un $\varepsilon > 0$ tel que la boule $V_\varepsilon(p)$ et l'ensemble B sont disjoints. D'après le lemme a, il existe un ensemble $A \in Q$ contenu entièrement dans la boule $V_\varepsilon(p)$. Évidemment, cet ensemble A n'a pas de points communs avec l'ensemble B , et nous aboutissons de nouveau à une contradiction. Le lemme est démontré.

d. On se sert souvent d'un cas particulier du lemme c que voici :

Lemme sur les boules fermées. *Dans un espace métrique complet, une suite des boules fermées emboîtées*

$$U_v = \{y : \rho(y, y_v) \leq r_v\}, \quad v = 1, 2, \dots,$$

dont les rayons r_v tendent vers zéro pour $v \rightarrow \infty$ d'une façon monotone, possède un point commun.

e. **Remarque.** S'il y a une suite d'intervalles fermés emboîtés de la droite numérique, ils possèdent toujours un point commun, indépendamment de leurs longueurs. Dans un espace métrique, même complet, il peut exister une suite de boules emboîtées n'ayant aucun point commun. Considérons, par exemple, l'espace formé d'une suite dénombrable des points $x_1, x_2, \dots$ munie de la distance définie par la formule $\rho(x_n, x_{n+p}) = 1 + \frac{1}{n}$ ($n = 1, 2, \dots$). Cet espace satisfait à tous les axiomes d'espace métrique. Il est complet puisqu'il n'admet aucune suite de Cauchy divergente (et même aucune suite de Cauchy composée de points distincts).

Une boule V_n de rayon $1 + \frac{1}{n}$ centrée en x_n contient les seuls points $x_n, x_{n+1}, \dots$. Il est évident que $V_1 \supset V_2 \supset \dots$ et que $\bigcap_{i=1}^{\infty} V_n$ est vide.

3.75.a. Théorème (Baire). *Si un espace métrique complet M est mis sous la forme d'une somme dénombrable de ses sous-ensembles fermés $A_1, A_2, \dots$, alors l'un au moins de ces sous-ensembles contient une boule de l'espace M .*

Démonstration. Supposons le contraire: aucun des ensembles $A_1, A_2, \dots$ ne contient une boule. Soit x_1 un point n'appartenant pas à l'ensemble A_1 ; celui-ci étant fermé, il existe une boule $V_{r_1}(x_1) = \{x : \rho(x, x_1) \leq r_1\}$, ne contenant aucun point de l'ensemble A_1 . À l'intérieur de la boule $V_{r_1/2}(x_1)$, il y a un point x_2 n'appartenant pas à l'ensemble A_2 ; en plus de x_2 , c'est une boule

$V_{r_2}(x_2)$ qui n'est pas contenue dans l'ensemble A_2 . On peut admettre que $V_{r_1}(x_1) \supset V_{r_2}(x_2)$ et $r_2 < \frac{r_1}{2}$. En continuant la construction de la même manière nous obtenons une suite des boules $V_{r_1}(x_1) \supset V_{r_2}(x_2) \supset \dots$, de sorte que la boule $V_{r_n}(x_n)$ n'a pas de points communs avec l'ensemble A_n ($n = 1, 2, \dots$). Le point x_0 commun à toutes ces boules, qui existe d'après le lemme 3.74.d, n'appartient à aucun des ensembles $A_1, A_2, \dots$, ce qui contredit la condition $x_0 \in M = \bigcup_{n=1}^{\infty} A_n$. Le théorème est démontré.

b. Exemple. Montrons que l'ensemble Z de tous les nombres irrationnels de l'intervalle $M = [a, b]$ ne peut être représenté sous forme de somme dénombrable de sous-ensembles fermés de M . En effet, si l'on avait $Z = \bigcup_{n=1}^{\infty} A_n$, où $A_n \subset M$ sont des ensembles fermés, alors tout l'intervalle M serait mis sous la forme d'une somme dénombrable de sous-ensembles fermés (d'une quantité dénombrable des ensembles A_n et d'une quantité dénombrable des ensembles réduits chacun à un seul point rationnel). L'intervalle $M = [a, b]$ étant un espace complet (3.73.c), on aboutirait au résultat contredisant le théorème de Baire.

c. En appliquant le théorème des intervalles fermés emboîtés nous avons démontré dans 2.41 que l'ensemble des points de l'intervalle $[0, 1]$ n'est pas dénombrable. A présent nous sommes en mesure d'établir une proposition analogue pour une vaste classe d'espaces métriques complets.

Donnons préalablement la définition suivante: un point x_0 d'un espace métrique M est dit *isolé* si une boule $p(x, x_0) < \delta$ ne contient pas de points de M autres que x_0 lui-même. Par exemple, soit M un ensemble de points de l'axe $-\infty < x < \infty$ avec sa métrique usuelle; alors $x_0 \in M$ est un point isolé s'il existe un intervalle contenant le point x_0 et ne contenant aucun autre point de l'ensemble M .

L e m m e. *Un espace métrique complet M , s'il est composé d'un ensemble dénombrable de points, contient un point isolé.*

D é m o n s t r a t i o n. Tout point est un sous-ensemble fermé d'un espace métrique. En appliquant le théorème de Baire nous voyons que, dans le cas considéré, un certain point x_0 de l'espace M contient une boule $V_r(x_0)$; ce n'est possible que lorsque le point x_0 est isolé.

Il en résulte:

d. Tout espace métrique complet M sans points isolés est non dénombrable.

Remarque. Cette proposition tombe en défaut si l'on rejette l'hypothèse sur l'absence de points isolés dans M . Tout ensemble dénombrable fermé de la droite considéré comme espace métrique peut en servir d'exemple (par exemple, une suite convergente et sa limite).

§ 3.8. Complétion

3.81. Les théorèmes démontrés dans 3.73-3.75 se basent essentiellement sur le fait que l'espace métrique M est complet. Dans ce qui suit, la complétude d'espaces métriques jouera un rôle important. D'autre part, il s'avère qu'un espace qui n'est pas complet peut être inclus dans un espace complet.

Théorème (F. Hausdorff). Soit M un espace métrique (en général, non complet). Il existe un espace métrique complet $\overline{M}$ (appelé *complété de l'espace M*) qui possède les propriétés suivantes:

- 1) M est isométrique à une partie $M_1 \subset \overline{M}$;
- 2) M_1 est dense dans $\overline{M}$.

N'importe quels deux espaces $\overline{M}, \overline{\overline{M}}$ vérifiant les conditions 1) et 2) sont isométriques l'un à l'autre.

La démonstration est donnée dans 3.82-3.87.

3.82. Deux suites de Cauchy $\{y_v\}$ et $\{z_v\}$ de l'espace M sont dites *cofinales* si $\lim_{v \rightarrow \infty} \rho(y_v, z_v) = 0$. Par exemple, toutes deux suites de l'espace M qui convergent vers une même limite sont cofinales, et les suites convergentes vers des limites différentes ne le sont pas. Deux suites de Cauchy cofinales chacune à une troisième sont cofinales entre elles. Pour cette raison, toutes les suites de Cauchy que l'on peut construire avec les éléments de l'espace M peuvent être partagées en classes de sorte que toutes les suites d'une classe sont cofinales entre elles, tandis que toute suite qui n'appartient pas à une classe n'est cofinale à aucune suite de cette classe. C'est de ces classes (on les désigne par $Y, Z, \dots$) que nous construirons le nouvel espace $\overline{M}$. Nous n'avons à définir que la valeur de la distance entre deux classes Y et Z . Nous la définissons avec la formule

$$\rho(Y, Z) = \lim_{v \rightarrow \infty} \rho(y_v, z_v), \quad (1)$$

où $\{y_v\}$ est une suite de Cauchy quelconque de la classe Y et $\{z_v\}$ une suite de Cauchy quelconque de la classe Z . Avant tout, il faut vérifier que la limite en question existe et ne dépend pas du choix des suites $\{y_v\}$ et $\{z_v\}$ dans les classes Y et Z respectivement. Il suit

de l'inégalité de quadrilatère (3.11 (1)) que

$$|\rho(y_v, z_v) - \rho(y_{v+\mu}, z_{v+\mu})| \leq \rho(y_v, y_{v+\mu}) + \rho(z_v, z_{v+\mu}).$$

Par conséquent, les nombres $\rho(y_v, z_v)$ forment une suite vérifiant le critère de Cauchy. Donc, $\lim_{v \rightarrow \infty} \rho(y_v, z_v)$ existe. Si $\{y'_v\}$ et $\{z'_v\}$ sont encore deux suites de Cauchy des classes Y et Z respectivement, alors en appliquant toujours l'inégalité de quadrilatère nous trouvons :

$$|\rho(y_v, z_v) - \rho(y'_v, z'_v)| \leq \rho(y_v, y'_v) + \rho(z_v, z'_v) \rightarrow 0,$$

ce qui signifie que la suite $\rho(y'_v, z'_v)$ a la même limite que la suite $\rho(y_v, z_v)$. Ainsi, la définition de la distance entre deux classes ne dépend pas du choix des suites de Cauchy dans ces classes.

3.83. A présent, nous avons à prouver que la quantité

$$\rho(Y, Z) = \lim_{v \rightarrow \infty} \rho(y_v, z_v)$$

satisfait aux axiomes de l'espace métrique (3.11.a-c).

L'axiome 3.11.a: $\rho(Y, Z) > 0$ pour $Y \neq Z$, $\rho(Y, Y) = 0$, est vérifié de la façon suivante. Avant tout, nous avons $\rho(Y, Y) = 0$ par construction de la fonction ρ , puisque, dans la formule 3.82 (1), on peut poser $y_v = z_v$. Ensuite, supposons que $\rho(Y, Z) = 0$. Ceci veut dire que, pour toute suite de Cauchy $\{y_v\}$ d'une classe Y et pour toute suite de Cauchy $\{z_v\}$ d'une classe Z , l'égalité $\lim_{v \rightarrow \infty} \rho(y_v, z_v) = 0$ a lieu. Or, ceci étant, $\{y_v\}$ et $\{z_v\}$ sont des suites cofinales, et les classes Y et Z doivent se confondre. Donc, si $\rho(Y, Z) = 0$, alors $Y = Z$; il en résulte que, pour $Y \neq Z$, nous avons $\rho(Y, Z) > 0$, ce qu'il fallait démontrer.

L'axiome 3.11.b: $\rho(Y, Z) = \rho(Z, Y)$, est satisfait par construction.

Vérifions l'axiome 3.11.c: $\rho(Y, U) \leq \rho(Y, Z) + \rho(Z, U)$. Soient $\{y_v\}$, $\{z_v\}$, $\{u_v\}$ des suites de Cauchy fixes des classes Y, Z, U respectivement. L'inégalité cherchée s'obtient en passant à la limite (3.32.c-d) dans l'inégalité

$$\rho(y_v, u_v) \leq \rho(y_v, z_v) + \rho(z_v, u_v).$$

Prouvons maintenant, pour l'espace $\bar{M}$, toutes les conclusions du théorème 3.81.

3.84. Montrons que $\bar{M}$ contient un sous-ensemble M , isométrique à l'espace M . Faisons correspondre à tout élément $y \in M$ la classe $Y \subset \bar{M}$ contenant la suite $y, y, y, \dots$ (i.e. la classe de toutes les suites convergentes vers y). Si, d'après cette loi, un point y correspond à une classe Y et un point z à une classe Z , alors

$$\rho(Y, Z) = \lim \rho(y, z) = \rho(y, z).$$

Il en résulte que l'ensemble M_1 des classes qui correspondent aux points de M est une partie de l'espace $\bar{M}$ isométrique à l'espace M .

3.85. Prouvons que M_1 est dense dans $\bar{M}$. Soient Y une classe quelconque de $\bar{M}$ et $\{y_v\}$ une suite de Cauchy de la classe Y . Considérons la suite des classes $Y_1, Y_2, \dots, Y_\mu, \dots$, où Y_μ est déterminée par la suite $y_\mu, y_\mu, \dots$, c'est-à-dire qu'elle correspond à l'élément y_μ selon la loi $M \rightarrow M_1$. Pour un $\varepsilon > 0$ donné, trouvons un numéro μ_0 de façon à avoir $\rho(y_{\mu+p}, y_\mu) \leq \varepsilon$ pour $\mu > \mu_0$. Alors

$$\rho(Y, Y_\mu) = \lim_{v \rightarrow \infty} \rho(y_v, y_\mu) \leq \varepsilon,$$

ce qui signifie que la classe Y est la limite des classes Y_μ . Toute classe Y_μ appartenant à l'ensemble M_1 par construction, il est démontré que M_1 est dense dans $\bar{M}$.

3.86. Montrons que $\bar{M}$ est un espace complet. Soit $Y_1, Y_2, \dots$ une suite de Cauchy d'éléments de $\bar{M}$. Pour toute classe Y_v , choisissons une classe $Z_v \subset M_1$ de façon à avoir $\rho(Y_v, Z_v) < \frac{1}{v}$, et soit $z_v \in M$ un élément correspondant à la classe Z_v . Démontrons que la suite $\{z_v\}$ est celle de Cauchy dans l'espace M . En effet,

$$\begin{aligned} \rho(z_v, z_\mu) &= \rho(Z_v, Z_\mu) \leq \rho(Z_v, Y_v) + \rho(Y_v, Y_\mu) + \\ &+ \rho(Y_\mu, Z_\mu) \leq \rho(Y_v, Y_\mu) + \frac{1}{v} + \frac{1}{\mu} \rightarrow 0 \text{ pour } v \rightarrow \infty, \mu \rightarrow \infty. \end{aligned}$$

La suite de Cauchy $\{z_v\}$ définit une classe $Z \subset \bar{M}$; montrons que la classe Z est la limite dans $\bar{M}$ de la suite Y_v . Pour un $\varepsilon > 0$ donné et pour un assez grand $v \geq v_0$, nous avons

$$\rho(Z, Y_v) \leq \rho(Z, Z_v) + \rho(Z_v, Y_v) = \lim_{\mu \rightarrow \infty} \rho(z_\mu, z_v) + \frac{1}{v} < \varepsilon.$$

Donc, toute suite de Cauchy $Y_v \subset \bar{M}$ a une limite dans $\bar{M}$, ce qu'il fallait démontrer.

3.87. Montrons, enfin, que tout espace métrique $\bar{M}$ possédant les propriétés 1) et 2) est isométrique à l'espace $\bar{M}$.

En effet, soient M_1 et M_2 des sous-ensembles des espaces $\bar{M}$ et $\bar{M}$, isométriques à l'espace M et, par conséquent, l'un à l'autre. Nous avons à prolonger cette isométrie à partir des ensembles M_1 et M_2 sur les espaces $\bar{M}$ et $\bar{M}$. Soient $Y \subset \bar{M}$ un élément quelconque et $Y_v \subset M_1$ une suite d'éléments convergente vers Y . La suite correspondante $Z_v \subset M_2$ est celle de Cauchy parce que, M_1 étant isométrique à M_2 , les distances entre les éléments de la suite Z_v sont égales à celles entre les éléments correspondants de la suite Y_v .

$\bar{M}$ étant complet, il existe dans $\bar{M}$ un élément $Z = \lim_{Y \rightarrow \infty} Z_Y$ que nous faisons correspondre à l'élément choisi plus haut $Y \subset \bar{M}$. Il est défini d'une façon univoque puisque les suites cofinales dans $\bar{M}$ correspondent aux suites cofinales dans $\bar{M}$ et en remplaçant la suite Y_Y par une suite cofinale on remplace la suite Z_Y par une suite cofinale. La correspondance construite est biunivoque et épuise tous les éléments de $\bar{M}$ et $\bar{M}$. Il nous reste à montrer qu'elle est isométrique. Supposons que deux éléments Y et Y' de l'espace $\bar{M}$ correspondent aux éléments Z et Z' respectivement de l'espace $\bar{M}$, et que

$$Y = \lim Y_Y,$$

$$Y' = \lim Y'_Y \quad (Y_Y, Y'_Y \text{ de } M_1).$$

Si, ensuite, Z_Y et Z'_Y sont deux éléments de M_2 correspondant aux éléments Y_Y et Y'_Y respectivement, alors $\rho(Z_Y, Z'_Y) = \rho(Y_Y, Y'_Y)$ et, d'après le lemme sur la continuité de la distance 3.32.b, on a

$$\rho(Z, Z') = \lim_{Y \rightarrow \infty} \rho(Z_Y, Z'_Y) = \lim_{Y \rightarrow \infty} \rho(Y_Y, Y'_Y) = \rho(Y, Y'),$$

ce qu'il nous fallait. Ainsi, le théorème 3.81 est complètement démontré.

3.88. Supposons qu'un espace métrique donné M soit une partie d'un autre espace métrique complet M^* . Alors on peut choisir pour complété de M la fermeture $\bar{M}$ de l'ensemble M dans l'espace M^* . En effet, $\bar{M}$ est un espace complet, puisqu'il est un sous-ensemble fermé d'un espace complet M^* (3.73.b); en outre, M est son sous-ensemble dense. $\bar{M}$ vérifie donc l'hypothèse du théorème démontré et peut servir de complété de l'espace M .

§ 3.9. Compacité

3.91.a. Définition. Un espace métrique M dans lequel toute suite (infinie) de points possède une valeur d'adhérence est appelé *espace compact* ou *compact*. Un espace métrique M dans lequel, pour tout point $a \in M$, il existe une boule compacte $\{x \in M : \rho(a, x) \leq c\}$ est appelé *espace localement compact*.

Si un espace métrique M est compact, tout sous-ensemble infini $E \subset M$ possède un point d'accumulation. En effet, on peut extraire d'un sous-ensemble infini dans E une suite infinie $x_1, x_2, \dots$ de points distincts dont la valeur d'adhérence est un point d'accumulation de l'ensemble E . Réciproquement, supposons que tout sous-ensemble infini de l'espace métrique M possède un point d'accumu-

lation; montrons que M est compact. Soit $x_1, x_2, \dots$ une suite de points de M (pas nécessairement distincts). Si cette suite ne comprend qu'un nombre fini de points distincts de l'espace M , alors au moins l'un d'eux se répète un nombre infini de fois et est donc une valeur d'adhérence de la suite $x_1, x_2, \dots$. Si elle comprend une infinité de points distincts de M , alors le point d'accumulation de cet ensemble infini est une valeur d'adhérence de toute la suite. Ainsi, la définition d'un espace compact donnée plus haut est équivalente à la définition suivante: *un espace métrique est compact si tout son sous-ensemble infini contient un point d'accumulation.*

b. Exemples. Un intervalle $a \leq x \leq b$ de la droite numérique est un compact (3.45). Toute la droite R n'est pas un espace compact, puisque la suite $1, 2, \dots, n, \dots$, par exemple, n'a pas de valeur d'adhérence dans R . Tout de même, R est un espace localement compact. La droite numérique achevée $\bar{R}$ munie de la métrique $r(x, y)$ (3.35.e) est compacte. L'ensemble des points rationnels d'un intervalle $[a, b]$ doté de la métrique de la droite numérique n'est pas compact ni localement compact.

c. Nous avons signalé dans 3.43 que la propriété d'un point donné d'être valeur d'adhérence d'une suite x_n n'est pas violée lorsqu'on passe à une métrique homéomorphe à la métrique initiale. Etant donné que les définitions d'un espace compact et d'un espace localement compact ne se basent que sur la notion de valeur d'adhérence, nous concluons que *les propriétés d'un espace d'être compact ou localement compact sont conservées si l'on passe à une nouvelle métrique homéomorphe à la métrique initiale.*

Ainsi, l'axe réel R est un espace localement compact pour sa métrique usuelle $\rho(x, y) = |x - y|$ aussi bien que pour la métrique $r(x, y)$ (3.35.e).

3.92.a. Théorème. *Tout espace compact est complet.*

Démonstration. Soient $x_1, x_2, \dots$ une suite de Cauchy de points d'un espace compact M et x_0 une valeur d'adhérence de cette suite. Montrons que $x_0 = \lim_{n \rightarrow \infty} x_n$. Pour un $\varepsilon > 0$ donné, trouvons d'abord un numéro N de façon à avoir $\rho(x_n, x_m) < \varepsilon/2$ pour $m > N, n > N$, puis un numéro $p > N$ tel que $\rho(x_p, x_0) < \varepsilon/2$. Alors, pour tous les $n > N$,

$$\rho(x_n, x_0) \leq \rho(x_n, x_p) + \rho(x_p, x_0) < \varepsilon,$$

d'où notre proposition.

b. Théorème. *Un sous-ensemble compact M d'un espace métrique P est fermé dans P .*

Ceci résulte de a et de 3.73.a.

c. Soit M un sous-ensemble compact d'un espace métrique P et soit $G \subset P$ un ensemble ouvert contenant M . Construisons l'ensemble ouvert M_δ qui est la réunion de toutes les boules ouvertes de rayon δ centrées en les points de l'ensemble M . Preuvons qu'il existe un $\delta > 0$ pour lequel $M_\delta \subset G$.

Raisonnons par l'absurde et trouvons, pour tout $n = 1, 2, \dots$, deux points $x_n \in M$, $y_n \in P - G$ pour lesquels $\rho(x_n, y_n) < \frac{1}{n}$. La suite x_n est dans le compact M , il est donc possible d'en extraire une sous-suite convergente; en numérotant d'une autre façon et en rejetant les points superflus on peut dire que $x_n \rightarrow x \in M$. Puisque $\rho(y_n, x) \leq \rho(y_n, x_n) + \rho(x_n, x) \leq \frac{1}{n} + \rho(x_n, x)$, alors $y_n \rightarrow x$. Or l'ensemble $P - G$ est fermé, donc $x \in P - G$ en contradiction avec ce qui précède. La proposition est démontrée.

3.93.a. Un élargissement de la classe des espaces compacts est représenté par la classe des espaces précompacts *). Un espace métrique M est dit *précompact* si toute suite de ses points contient une sous-suite de Cauchy. Si, de plus, M est complet, cette suite de Cauchy est convergente et l'espace M se trouve compact. Ainsi, un *espace précompact complet est compact*. Réciproquement, un *espace compact est complet* (3.92.a) et, évidemment, *précompact*. Un intervalle (a, b) de la droite numérique représente un exemple simple d'un espace précompact qui n'est pas compact.

b. Théorème. *Tout espace précompact est borné.*

Démonstration. Nous allons démontrer qu'un espace qui n'est pas borné n'est pas précompact. Dans un espace non borné M , les distances d'un point quelconque a aux autres points $x \in M$ ne sont pas majorées par une constante (3.12.c). En appliquant cette propriété fixons arbitrairement un point $x_1 \in M$ et construisons ensuite par récurrence une suite des points $x_2, x_3, \dots$ de manière à satisfaire aux inégalités

$$\rho(x_{n+1}, x_n) > \sum_{k=1}^{n-1} \rho(x_{k+1}, x_k) + 1 \quad (n = 1, 2, \dots).$$

Alors, d'après 3.11.d, pour $n > m$,

$$\begin{aligned} \rho(x_n, x_m) &\geq \rho(x_n, x_{n-1}) - [\rho(x_{n-1}, x_{n-2}) + \dots + \rho(x_{m+1}, x_m)] \geq \\ &\geq \rho(x_n, x_{n-1}) - \sum_{k=1}^{n-2} \rho(x_{k+1}, x_k) > 1, \end{aligned}$$

*) Cette terminologie n'est pas universellement admise. Les espaces précompacts sont parfois appelés espaces compacts et les espaces compacts les compacts tout court.

de sorte qu'il est impossible d'extraire de la suite x_n une sous-suite de Cauchy. L'espace M n'est donc pas précompact, ce qu'il fallait démontrer.

c. Critère de Hausdorff. Pour voir si un espace métrique M est précompact, il est parfois utile de l'inclure isométriquement dans un espace métrique plus large P . En supposant isométrique l'inclusion $M \subset P$ appelons un ensemble $B \subset P$ *e-réseau de l'ensemble $M \subset P$* si la distance de tout point x de l'ensemble M à un point $y \in B$ ne dépasse pas e . Donc, la réunion de toutes les boules de rayon e centrées en les points de l'ensemble B contient tout l'ensemble M .

D'une manière générale, si la réunion de quelques ensembles U_α contient un ensemble M , on dit que les ensembles U_α recouvrent l'ensemble M ou bien qu'ils forment un *recouvrement* de l'ensemble M . On peut donc dire qu'un ensemble B est un *e-réseau* pour un ensemble M si l'ensemble de toutes les boules de rayon e centrées en les points de l'ensemble B recouvre l'ensemble M *).

Théorème (F Hausdorff). *Un ensemble M contenu dans un espace métrique P est précompact (pour la métrique de P) s'il existe dans P , pour tout $e > 0$, un e-réseau fini pour M , et dans ce cas seulement.*

Démonstration. Soient M un ensemble précompact et un $e > 0$. Il existe un *e-réseau fini* pour M . Considérons un point quelconque $x_1 \in M$. Si tout autre point de l'ensemble M est situé à une distance $\leq e$ du point x_1 , alors le point x_1 représente lui-même un *e-réseau* pour M et la construction est terminée. Si M possède des points distants de x_1 plus que de e , nous en choisissons un quelconque x_2 . Si maintenant tout point de l'ensemble M est distant au plus de e du point x_1 ou du point x_2 , alors x_1 et x_2 forment un *e-réseau fini* pour M et la construction est terminée; sinon, on continue de la même manière. Par construction, la distance de x_n à chacun des points précédents $x_1, x_2, \dots, x_{n-1}$ est supérieure à e . Pour cette raison, s'il était possible de continuer indéfiniment la construction, on arriverait à un sous-ensemble infini $x_1, x_2, \dots, x_n, \dots$ de l'ensemble M qui ne contiendrait aucune suite de Cauchy, ce qui contredirait la précompacité de M . M étant précompact, ladite construction est terminée après un nombre fini de pas et conduit à un *e-réseau fini* pour l'ensemble M .

Réciproquement, supposons qu'il existe dans l'espace P , pour tout $e > 0$, un *e-réseau fini* pour l'ensemble M ; montrons que M est précompact. Soit $A \subset M$ un sous-ensemble infini quelconque;

*) « Si, en tout point d'un ensemble B qui est un *e-réseau* pour un ensemble M , on allume une lampe éclairant une boule de rayon e , alors tout l'ensemble M sera éclairé. » (Extrait d'un cours de L. Lusternik.)

nous avons à trouver dans A une suite de Cauchy. Pour premier point de cette suite nous choisissons un point quelconque $x_0 \in A$. En appliquant l'hypothèse du théorème avec $\varepsilon = 1$ nous sommes en mesure de recouvrir l'ensemble A par un nombre fini de boules de rayon 1 dont l'une (désignons-la par U_1) contient un sous-ensemble infini $A_1 \subset A$. Choisissons dans A_1 un point quelconque $x_1 \neq x_0$. En appliquant l'hypothèse du théorème avec $\varepsilon = \frac{1}{2}$ nous pouvons recouvrir l'ensemble A_1 par un nombre fini de boules de rayon $\frac{1}{2}$ dont l'une (U_2) contient un sous-ensemble infini $A_2 \subset A_1$. Choisissons un point $x_2 \in A_2$ différent de x_0 et de x_1 . En poursuivant nous construisons une chaîne d'ensembles infinis $A \supset A_1 \supset A_2 \supset \dots \supset A_v \supset \dots$ (tout ensemble A_v est contenu dans la boule U_v de rayon $\frac{1}{v}$) aussi bien qu'une suite de points distincts $x_0, x_1, x_2, \dots, x_v, \dots$, où $x_v \in A_v$. La suite $x_0, x_1, x_2, \dots$ est de Cauchy. En effet, pour $\mu < v$, nous avons $U_\mu \supset A_\mu \supset A_v$, donc $\rho(x_\mu, x_v) < \frac{2}{\mu}$. Cette quantité tend vers zéro pour $\mu \rightarrow \infty$; par conséquent, la suite $x_0, x_1, \dots$ est de Cauchy, ce qu'il fallait démontrer.

3.94. Appliquons ce critère pour démontrer que *tout ensemble borné M dans un espace euclidien n -dimensionnel $P = R_n$ est précompact*. En effet, pour tout m , une boule dans l'espace P qui contient l'ensemble borné M (3.14.b) n'a qu'un nombre fini de points dont les coordonnées sont de la forme $\frac{k}{2^m}$ avec k entier, et l'ensemble de tous les points de ce type est, évidemment, un ε -réseau pour M si m est assez grand.

3.95. Signalons encore une simple condition suffisante de précompactité : *un ensemble M est précompact dans un espace métrique P si, pour tout $\varepsilon > 0$, on peut trouver dans P un ensemble précompact B_ε (peut-être infini) représentant un ε -réseau pour M .*

La démonstration de cette proposition est très simple. Pour un ε donné, un $\frac{\varepsilon}{2}$ -réseau fini Z pour l'ensemble $B_{\varepsilon/2}$ qui existe en vertu de la précompactité de $B_{\varepsilon/2}$ est un ε -réseau pour l'ensemble M . En effet, pour un point quelconque $x \in M$, il existe par hypothèse un point $y \in B_{\varepsilon/2}$ tel que $\rho(x, y) \leq \frac{\varepsilon}{2}$ et un point $z \in Z$ tel que $\rho(y, z) \leq \frac{\varepsilon}{2}$; alors $\rho(x, z) \leq \rho(x, y) + \rho(y, z) \leq \varepsilon$. Donc, pour tout $\varepsilon > 0$, l'ensemble M possède un ε -réseau fini et, par conséquent, est précompact.

3.96.a. *Le complété $\overline{M}$ de tout ensemble précompact M est un compact*. En effet, puisque l'ensemble M est dense dans $\overline{M}$, il est

un ε -réseau pour $\bar{M}$, pour tout $\varepsilon > 0$. Par hypothèse, M est précompact; il en résulte, d'après 3.95, que $\bar{M}$ l'est aussi; or $\bar{M}$ est complet, donc il est un compact, ce que nous cherchions.

b. *La fermeture $\bar{M}$ de tout sous-ensemble précompact M d'un espace métrique complet P est compacte.* Ceci résulte de a et de la remarque 3.88 d'après laquelle on peut choisir pour complétude de M sa fermeture dans P .

c. *Un sous-ensemble précompact M d'un espace métrique complet P est un compact si, et seulement si, il est fermé dans P .*

En effet, si le sous-ensemble M est fermé dans l'espace métrique complet P , il est lui-même un espace métrique complet (3.73.b). Si, de plus, M est précompact, alors, d'après 3.93.a, il est un compact. La réciproque résulte de 3.92.b.

d. En rapprochant 3.93.c et 3.96.c, on obtient:

T h é o r è m e. *Un ensemble M dans un espace métrique complet P est un compact si, et seulement si, il est fermé dans P et que, pour tout $\varepsilon > 0$, il existe dans P un ε -réseau fini pour M .*

e. *Un ensemble M dans un espace R_n est un compact si, et seulement si, M est fermé et borné dans R_n .* En effet, si M est un compact, alors l'ensemble M est fermé (3.92.b) et borné (3.93.b); si l'ensemble M est borné dans R_n , il est précompact (3.94), donc compact (3.96.c), étant donné que R_n est complet (3.72.c). En particulier, une boule fermée $\{x: |x - x_0| \leq r\}$ est un ensemble fermé (3.51.b) et borné, donc toute boule fermée dans l'espace R_n est compacte.

f. *Tout ensemble compact M de la droite numérique est borné et contient ses bornes supérieure et inférieure.*

En effet, l'ensemble M est borné et fermé (e); l'application de 3.53.b achève la démonstration.

3.97. La proposition suivante concernant les compacts est souvent appliquée:

L e m m e s u r l e r e c o u v r e m e n t f i n i. *Si une famille $B = \{B_\alpha\}$ de sous-ensembles ouverts de l'espace P recouvre un sous-ensemble compact K d'un espace métrique P , alors il existe une sous-famille finie $B_1, \dots, B_m$ de B recouvrant K .*

D é m o n s t r a t i o n. Supposons le contraire: aucune sous-famille finie de la famille B ne recouvre K .

Comme K est un compact, il existe, pour tout $\varepsilon > 0$, d'après 3.93.c, un nombre fini de boules (fermées) $U_1, \dots, U_m$, de rayon ε qui recouvrent K . S'il existait, pour chacune des boules U_j ($j = 1, \dots, m$), une sous-famille finie B_j de la famille B recouvrant la boule U_j , alors en réunissant ces sous-familles nous aurions une sous-famille finie de B recouvrant K . Donc, l'une au moins des

boules U_j , par exemple U_1 , ne peut être recouverte par aucune sous-famille finie de la famille B .

Donnons au nombre ε des valeurs successives $1, \frac{1}{2}, \frac{1}{3}, \dots$; pour tout n , nous avons une boule $U_{1/n}(x_n)$ de rayon $\frac{1}{n}$ et de centre x_n qui ne peut être recouverte par aucune sous-famille finie de la famille B . Soit x une valeur d'adhérence de la suite x_n . Le point x appartient à un ensemble B_α avec une boule $U_\rho(x)$ d'un rayon ρ . A partir d'un certain numéro, la boule $U_\rho(x)$ contient les boules $U_{1/n}(x_n)$ avec les numéros arbitrairement grands. Donc, ces boules sont recouvertes par un seul ensemble B_α , ce qui contredit la construction. Le lemme est démontré.

3.98. Nous avons étendu (3.74) le principe d'intervalles fermés emboîtés (1.81) de la droite numérique à un espace métrique complet en remplaçant dans son énoncé les intervalles fermés par les sous-ensembles fermés de diamètres arbitrairement petits. Si ces sous-ensembles sont compacts, on n'a plus besoin d'hypothèses supplémentaires :

Théorème. *Toute famille Q de sous-ensembles emboîtés compacts non vides d'un espace métrique a un point commun.*

Démonstration. Fixons dans la famille Q l'un des compacts K_0 . Si le théorème est faux, il existe, pour tout point $x \in K_0$, un compact $K_x \in Q$ qui ne contient pas le point x . L'ensemble K_x étant fermé (3.92.b), il existe un voisinage du point x sans point commun avec K_x . Les voisinages ainsi obtenus pour tous les points $x \in K_0$ forment un recouvrement du compact K_0 ; d'après 3.97, on peut en dégager un recouvrement fini. Désignons les voisinages composant le recouvrement fini par $V_1, \dots, V_n$. Soient $K_1, \dots, K_n$ les compacts de la famille Q sans points communs respectivement avec $V_1, \dots, V_n$. L'intersection $K_1, \dots, K_n$ n'a de points communs avec aucun des voisinages $V_1, \dots, V_n$, ni donc avec le compact K_0 . Ainsi, l'intersection $K_0 K_1 \dots K_n$ est vide. Or, comme Q est une famille d'ensembles emboîtés, toute intersection finie d'ensembles de cette famille est encore un ensemble de cette famille et ne peut donc être vide. La contradiction obtenue démontre le théorème.

Exercices

1. Désignons par A' l'ensemble des points d'accumulation d'un sous-ensemble A d'un espace métrique M . Définissons ensuite, par récurrence, l'ensemble $A^{(n)} = (A^{(n-1)})'$. Pour un n donné, construis sur la droite un ensemble A tel que $A^{(n)}$ soit non vide et $A^{(n+1)}$ vide.

2. Démontrer que A' est fermé pour tout $A \subset M$.

3. On se donne un ensemble $A \subset R$, tel que, pour un certain n , l'ensemble $A^{(n)}$ est dénombrable. Démontrer que A est dénombrable.

4. Un point x de la droite est appelé *point de condensation* d'un ensemble non dénombrable A si tout son voisinage contient un ensemble non dénombrable de points de l'ensemble A . Démontrer que tout ensemble A non dénombrable a des points de condensation; plus exactement, presque tous ses points, à l'exception, peut-être, d'un ensemble dénombrable, sont des points de condensation.

5. Si un ensemble E de la droite est recouvert par un système quelconque d'intervalles ouverts, alors on peut en extraire un sous-système au plus dénombrable qui recouvre lui aussi E .

6. La quantité

$$\rho(x, A) = \inf_{y \in A} \rho(x, y)$$

s'appelle *distance d'un point x à un ensemble A* . Démontrer que, pour un ensemble A fermé, les relations

$$\omega(x, A) = 0, \quad x \in A$$

sont équivalentes et que pour un A qui n'est pas fermé, elles ne le sont pas.

7. Démontrer que, pour tout ensemble A dans un espace métrique M , l'ensemble des points x pour lesquels $\rho(x, A) < \varepsilon$ est ouvert, et l'ensemble des points y pour lesquels $\rho(y, A) \leq \varepsilon$ est fermé.

8. On se donne deux ensembles fermés disjoints F_1 et F_2 dans un espace métrique M . Construire deux ensembles ouverts disjoints U_1 et U_2 tels que $U_1 \supset F_1$, $U_2 \supset F_2$.

9. Soit M un espace métrique borné. Pour deux sous-ensembles quelconques $A \subset M$ et $B \subset M$, introduisons la quantité

$$\rho(A, B) = \sup_{\substack{x \in A \\ y \in B}} \rho(x, y).$$

Montrer que tous les sous-ensembles fermés $A, B, \dots$ dans M forment un espace métrique si l'on définit la distance entre deux sous-ensembles fermés par la formule donnée ci-dessus. Montrer que cet espace est complet si M est complet, et compact si M est compact.

10. Si un espace métrique M est composé de $n < 4$ points, il existe un espace métrique M' isométrique à M et situé dans l'espace euclidien R_{n-1} . Pour $n \geq 4$, cette proposition est, en général, fausse.

11. Désignons par $\bar{R}_n$ l'espace euclidien R_n complété par le point ∞ . Définir sur $\bar{R}_n$ une métrique r qui soit homéomorphe sur R_n à la métrique usuelle ρ (3.14) et telle que toute suite $x_m \in R_n$ non bornée pour la métrique usuelle ait le point ∞ pour sa valeur d'adhérence.

12. Faire l'exercice 11 en remplaçant l'espace R_n par un espace métrique M non borné quelconque.

13. Dans l'exercice 11, prendre au lieu des droites considérées (cf. l'indication pour l'exercice 11) les droites passant par le centre de la sphère S_n . Quel ensemble d'éléments « à l'infini » assurera l'existence d'une valeur d'adhérence de toute suite de points $x_m \in R_n$ (pour la nouvelle métrique)?

Historique

La première définition correcte de la limite d'une suite numérique est donnée par Bolzano (1817), puis par Cauchy (1821) dans son *Cours d'analyse de l'école polytechnique*. En particulier, Bolzano est le premier à formuler en termes clairs le « critère de Cauchy » et même à essayer de l'établir; mais son raisonne-

mont, « en l'absence de toute définition arithmétique du nombre réel, n'était et ne pouvait être qu'un cercle vicieux » (Bourbaki). Cauchy lui-même tire son critère du principe d'intervalles fermés emboîtés qu'il tient pour évident.

La notion de point d'accumulation d'un ensemble ouvert ou fermé (sur la droite d'abord, dans l'espace euclidien n -dimensionnel ensuite) et les théorèmes sur la structure de ces ensembles pour la droite sont donnés par Cantor dans les années 70 du XIX^e siècle. Le théorème sur l'extraction d'un recouvrement fini est démontré pour la première fois (pour un intervalle fermé) par Borel (1895; cas où le recouvrement initial est dénombrable) et Lebesgue (1902; cas général). En 1906, Fréchet introduit la notion d'espace métrique dans le cadre de laquelle les notions de la période précédente sont généralisées d'une façon naturelle; en outre, Fréchet introduit les notions de complétude et de compacité d'un espace métrique.

La notion d'espace topologique (Hausdorff, 1914) que nous ne citons ici qu'en passant offre des possibilités encore plus vastes [19].

Théorie générale des limites

L'œuvre excellente de M. Cauchy
« Cours d'Analyse de l'Ecole Royale
Polytechnique » que chaque analyste ai-
mant la rigueur dans la recherche mathé-
matique doit lire se servait de guide.

Niels Abel (1826)

§ 4.1. Définition d'une limite

4.11. Définition. Soit un ensemble quelconque E ; un système S de sous-ensembles non vides $A, B, \dots$ de E s'appelle *direction* si, pour tout couple de sous-ensembles $A \in S, B \in S$, l'une des inclusions $A \subset B$ et $B \subset A$ a lieu et l'intersection de tous les $A \in S$ est vide.

4.12. Définition. Soit $y = f(x)$ une fonction donnée sur un ensemble E et dont les valeurs appartiennent à un espace métrique M ayant ρ pour métrique (3.11). Nous dirons que la fonction $y = f(x)$ a une limite suivant une direction S s'il existe un point $p \in M$ tel que, pour tout $\epsilon > 0$, on peut trouver un ensemble $A \in S$ pour tous les points duquel on a l'inégalité

$$\rho[p, f(x)] < \epsilon. \quad (1)$$

Dans ce cas, le point p s'appelle *limite de la fonction $f(x)$ suivant la direction S* . Toute la situation décrite est désignée par le symbole

$$p = \lim_S f(x), \quad (2)$$

ou $f(x) \xrightarrow[S]{} p$, ou tout simplement $f(x) \rightarrow p$.

Dans 4.13-4.16, on considère des exemples.

4.13. Soit E l'ensemble de tous les nombres naturels $1, 2, \dots$. Choisissons pour direction S le système de tous les sous-ensembles $A_n \subset E$ de la forme

$$A_n = \{n, n+1, n+2, \dots\} \quad (n = 1, 2, \dots).$$

Il est évident que l'un de deux ensembles quelconques A_n, A_m contient l'autre; de plus, l'intersection de tous les ensembles A_n ($n = 1, 2, \dots$) est vide. Donc, le système S est bien une direction

que l'on désigne par $n \rightarrow \infty$. La fonction $y = f(x)$ est, dans ce cas, une suite de points $y_1, y_2, \dots, y_n, \dots$ de l'espace métrique M . D'après notre définition, la suite y_n a une limite suivant S , i.e. pour $n \rightarrow \infty$, s'il existe un point $p \in M$ tel que, pour tout $\varepsilon > 0$, on peut trouver un numéro N de façon que l'inégalité

$$\rho(y_n, p) \leq \varepsilon$$

soit satisfaite pour tout $n \geq N$. Une suite $\{y_n\}$ possédant une limite p est dite *convergente vers p* . Il est évident que cette définition se confond avec celle de 3.31. La désignation 4.12 (2) prend la forme $p = \lim_{n \rightarrow \infty} y_n$.

4.14.a. Soit $E = R_a^+$ un demi-axe réel $\{x: x \geq a\}$. Choisissons pour direction S le système de tous les sous-ensembles $A_\xi \subset R_a^+$ de la forme

$$A_\xi = \{x \in R_a^+ : x \geq \xi\}.$$

Il est évident que S est une direction au sens de 4.11. On la note $x \rightarrow +\infty$. D'après notre définition, une fonction $f(x)$ définie pour $x \geq a$ (à valeurs dans un espace métrique M) a une limite suivant S , i.e. pour $x \rightarrow +\infty$, s'il existe un point $p \in M$ tel que, pour tout $\varepsilon > 0$, on peut trouver un nombre ξ de façon à avoir, pour tout $x \geq \xi$, l'inégalité

$$\rho(p, f(x)) \leq \varepsilon.$$

Le point p s'appelle *limite de la fonction $f(x)$ pour $x \rightarrow \infty$* ; la notation respective a la forme

$$p = \lim_{x \rightarrow \infty} f(x).$$

b. Si $M = R$ est la droite numérique, nous obtenons la définition d'une fonction numérique ayant une limite pour $x \rightarrow \infty$, notamment : une fonction numérique $f(x)$ définie pour $x \geq a$ a une limite p , lorsque $x \rightarrow \infty$, si, pour tout $\varepsilon > 0$, il existe un point ξ tel que l'inégalité

$$|p - f(x)| < \varepsilon$$

est satisfaite, dès que $x \geq \xi$.

c. Sur un demi-axe réel $R_a^- = \{x: x \leq a\}$, on introduit la direction $x \rightarrow -\infty$. Elle est définie comme système de tous les sous-ensembles $B_\xi \subset R_a^-$ de la forme

$$B_\xi = \{x \in R_a^- : x \leq \xi\} = R_\xi^-.$$

D'après notre définition, une fonction $f(x)$ définie pour $x \leq a$ a une limite p , pour $x \rightarrow -\infty$, si, quel que soit un $\varepsilon > 0$, on peut

trouver un nombre ξ tel que, pour tout $x \leq \xi$, on a l'inégalité

$$\rho(p, f(x)) \leq \varepsilon$$

ou bien, si la fonction $f(x)$ est numérique,

$$|p - f(x)| \leq \varepsilon.$$

d. D'une manière analogue, on introduit la direction $x \rightarrow \pm \infty$, ou $|x| \rightarrow \infty$, sur la droite numérique. La direction $|x| \rightarrow \infty$ est formée de tous les ensembles $\{x: |x| \geq C\}$. Ainsi, l'expression

$$\lim_{|x| \rightarrow \infty} f(x)$$

prend un sens.

4.15.a. Soit E un espace métrique. Supposons qu'un point $a \in E$ ne soit pas un point isolé dans E , autrement dit, que toute boule $U_r(a) = \{x \in E: \rho(x, a) < r\}$ contienne des points de E autres que a . Définissons la direction $x \rightarrow a$ comme ensemble de toutes les boules $U_r(a) = \{x \in E: \rho(a, x) < r\}$ dépourvues du point central a . La supposition que a ne soit pas un point isolé veut dire qu'aucun des ensembles $U_r(a)$ ne soit vide; les autres propriétés de la direction sont immédiates. D'après notre définition, une fonction $f(x)$ définie sur E (à valeurs dans un espace métrique M) a une limite pour $x \rightarrow a$ s'il existe un point $p \in M$ tel que, quel que soit un $\varepsilon > 0$, on peut trouver un nombre $\delta > 0$ de façon à avoir

$$\rho(f(x), p) \leq \varepsilon$$

pour tous les $x \in U_\delta(a)$ (i.e. pour tous les $x \neq a$ satisfaisant à l'inégalité $\rho(x, a) < \delta$).

La notation 4.12 (2) prend la forme

$$p = \lim_{x \rightarrow a} f(x). \quad (1)$$

Le point p s'appelle toujours *limite de la fonction $f(x)$ pour $x \rightarrow a$* *).

b. En particulier, lorsque $E = R$, $M = R$, nous sommes amenés à la définition suivante:

La direction $x \rightarrow a$ est l'ensemble de tous les intervalles $(a - \delta, a + \delta)$ dépourvus du point central a . Une fonction numérique $f(x)$ définie sur la droite numérique R possède une limite p au point $a \in R$ si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que les relations $|x - a| < \delta$, $x \in R$, $x \neq a$ impliquent l'inégalité

$$|f(x) - p| \leq \varepsilon.$$

*) Dans cette définition, la valeur de la fonction $f(x)$ pour $x = a$ est sans importance. La fonction $f(x)$ peut même ne pas être définie pour $x = a$.

La notation (1) ne change pas :

$$p = \lim_{x \rightarrow a} f(x).$$

c. Les exemples 4.13 et 4.14.a-c sont bien des cas particuliers de la définition 4.15.a. Vérifions-le pour les exemples 4.14.a-c. On peut munir la droite numérique R de la métrique de l'espace $\bar{R}$ (3.35.e); alors, les directions $x \rightarrow -\infty$ et $x \rightarrow +\infty$ décrites dans 4.14.a se confondent avec les directions $x \rightarrow -\infty$ et $x \rightarrow +\infty$ respectivement (4.15.a), où $-\infty$ et $+\infty$ sont considérés en tant que points de l'espace métrique $\bar{R}$ (compte tenu de 3.35.f).

4.16. Limite suivant une direction sur un sous-ensemble.

a. Soit E un ensemble contenant une direction S . Fixons un ensemble $G \subset E$ et considérons la famille des ensembles GA , où A est un ensemble quelconque de la direction S . Supposons qu'aucun des ensembles GA ne soit vide. Alors, l'intersection de ces ensembles étant vide (puisque l'intersection de tous les $A \in S$ l'est), ils forment encore une direction que nous désignons par GS . Soit $f(x)$ une fonction définie sur l'ensemble E à valeurs dans un espace métrique M . Si $\lim_{S} f(x)$ valant p existe, il est évident que $\lim_{GS} f(x)$ qui vaut p existe elle aussi. Mais, étant donné que $\lim_{GS} f(x)$ existe, $\lim_{S} f(x)$ peut ne pas exister. Établissons la proposition suivante qui fournit une condition d'équivalence de deux limites en question :

L e m m e. Si l'ensemble G contient un ensemble B de la direction S , alors l'existence de $\lim_{GS} f(x)$ implique l'existence de $\lim_{S} f(x)$ ($= \lim_{GS} f(x)$). Si l'ensemble G ne contient entièrement aucun $B \in S$ et que l'espace métrique M ait au moins deux points distincts p et $q \neq p$, alors il existe une fonction $f(x)$ à valeurs dans M pour laquelle $\lim_{GS} f(x) = p$ tandis que $\lim_{S} f(x)$ n'existe pas.

D é m o n s t r a t i o n. Soient $G \supset B$, $B \in S$ et $\lim_{GS} f(x) = p$. Pour un $\varepsilon > 0$ donné, choisissons, dans la direction GS , un ensemble GA dont chaque point vérifie l'inégalité $\rho(f(x), p) < \varepsilon$. L'ensemble GA contient l'ensemble BA qui est ou bien B ou bien A , donc appartient à la direction S . L'inégalité $\rho(f(x), p) < \varepsilon$ est aussi satisfaite sur l'ensemble BA . Il en résulte que $\lim_{S} f(x)$ existe (et vaut p).

Supposons que l'ensemble G ne contienne entièrement aucun $B \in S$, et soit H le complémentaire de G (par rapport à E). Posons $f(x)$ égale à p pour $x \in G$ et à q pour $x \in H$. Soit $\varepsilon < \frac{1}{2} \rho(p, q)$.

Si $\lim_S f(x) = t$ existait, alors on aurait $\rho(f(x), t) < \varepsilon$ pour un $A \in S$ et pour tout $x \in A$. Or, les deux ensembles GA et HA sont non vides par hypothèse; considérant, à tour de rôle, $x \in GA$ et $x \in HA$ nous voyons que les deux inégalités $\rho(p, t) < \varepsilon$, $\rho(q, t) < \varepsilon$ devraient être satisfaites, d'où $\rho(p, q) \leq \rho(p, t) + \rho(q, t) < 2\varepsilon$ en contradiction avec la définition de ε . Le lemme est donc définitivement démontré.

b. Si une fonction $f(x)$ n'est définie que sur l'ensemble G , la notation $\lim_S f(x)$ n'a pas de sens direct. Dans le cas où G contient un ensemble de la direction S , nous posons par définition $\lim_S f(x) = \lim_S f_E(x)$, où $f_E(x)$ est un prolongement quelconque de la fonction $f(x)$ de G sur tout E . D'après ce qu'on a démontré, le résultat a un sens et ne dépend pas de la manière de prolongement de $f(x)$ de G sur E .

En particulier, la limite d'une suite $y_1, y_2, \dots, y_n, \dots$ (4.13) est, naturellement, définie non seulement dans le cas où la suite $y_1, y_2, \dots$ est donnée pour tous les indices $1, 2, \dots$, mais aussi lorsqu'elle n'est donnée que pour tous les indices n supérieurs à un n_0 ; alors $\lim_{n \rightarrow \infty} y_n$ ne dépend pas des valeurs $y_1, \dots, y_{n_0}$ si nous intro-

duisons ces dernières pour une raison quelconque. D'une façon analogue, la définition de $\lim_{n \rightarrow \infty} f(x)$ (4.14) suppose que seules les

valeurs de la fonction $f(x)$ pour x supérieurs à un x_0 quelconque soient données, et ne dépend pas des valeurs de $f(x)$ pour $x \leq x_0$. De même, la définition de $\lim_{x \rightarrow a} f(x)$ (4.15) ne demande que de donner les va-

leurs de la fonction $f(x)$ dans un voisinage du point $x = a$, c'est-à-dire dans une boule $\rho(x, a) < r$, $x \neq a$.

c. Revenons au cas où l'ensemble G , tout en ayant des points communs avec chacun des ensembles $A \in S$, ne contient entièrement aucun $A \in S$. Comme plus haut, désignons par H le complémentaire de G par rapport à E . Considérons la famille HS des ensembles HA , où A parcourt S . Puisque, dans ce cas, aucun HA n'est vide, le système HS forme lui aussi une direction, et nous pouvons discuter l'existence de $\lim_{HS} f(x)$. Si $\lim_S f(x) = p$ existe, alors $\lim_{GS} f(x)$ et

$\lim_{HS} f(x)$ égales chacune à p existent aussi. Néanmoins, l'exemple

donné dans la seconde partie de la démonstration du lemme a montre que l'existence des $\lim_{GS} f(x)$ et $\lim_{HS} f(x)$ n'implique pas

l'existence de $\lim_S f(x)$.

Théorème. Si $\lim_{GS} f(x)$ et $\lim_{HS} f(x)$ existent et se confondent, $\lim_S f(x)$ existe aussi.

Démonstration. Soit $p = \lim_{GS} f(x) = \lim_{HS} f(x)$. Pour un $\varepsilon > 0$ donné, trouvons A et B dans S de façon que l'inégalité

$$\rho(f(x), p) < \varepsilon \quad (1)$$

soit satisfaite aux points des ensembles GA et HB .

L'un des deux ensembles A et B est inclus dans l'autre, par exemple $A \subset B$. Alors l'inégalité (1) est bien satisfaite pour tous les points de l'ensemble GA et de l'ensemble HA , donc aussi pour tout point de l'ensemble $A = GA + HA$. Comme $\varepsilon > 0$ est arbitraire, on a $\lim_S f(x) = p$, ce qu'on cherchait.

4.17. Applications bijectives de l'ensemble E et transformations correspondantes de limites.

a. Supposons qu'une application bijective de l'ensemble E sur un ensemble F fait correspondre à tout $x \in E$ un $y = \omega(x) \in F$. Soient S une direction sur l'ensemble E formée d'ensembles $A \subset E$ et $B \subset F$ l'image d'un ensemble A par l'application ω . L'ensemble T de tous les ensembles B forme une direction sur F parce qu'il résulte des propriétés de l'application bijective ω que les ensembles B , de même que les ensembles A , sont emboîtés et que leur intersection est vide. Soit, ensuite, $f(x)$ une fonction donnée sur l'ensemble E à valeurs dans un espace métrique M . Définissons une fonction $g(y)$ sur l'ensemble F par la formule

$$g(y) \equiv g(\omega(x)) = f(x). \quad (1)$$

Théorème. La fonction $g(y)$ a une limite suivant T si, et seulement si, la fonction $f(x)$ a une limite suivant S , et alors

$$\lim_S f(x) = \lim_T g(y).$$

Démonstration. Supposons que $\lim_T g(y) = p$ existe. Pour un $\varepsilon > 0$ donné, on peut trouver un ensemble $B \in T$ tel que $\rho(g(y), p) < \varepsilon$ pour $y \in B$. Sur l'ensemble correspondant $A \in S$, nous avons

$$\rho(f(x), p) = \rho(g(\omega(x)), p) < \varepsilon,$$

de sorte que $p = \lim_S f(x)$. La réciproque résulte de la symétrie de la construction, ce qui achève la démonstration.

b. Il découle de 4.16.b que le résultat précédent reste valable si l'application bijective ω est définie non pas sur tout le E , mais sur un quelconque des sous-ensembles $A \in S$.

Considérons deux exemples.

c. Choisissons pour E un demi-axe $\{x: x > a\}$ et soit $y = -x$ l'application du demi-axe E sur le demi-axe $F = \{y: y < -a\}$. Choisissons sur E la direction $x \rightarrow \infty$. La direction correspondante sur F est évidemment $y \rightarrow -\infty$. D'après a, les limites $\lim_{x \rightarrow \infty} f(x)$ et $\lim_{y \rightarrow -\infty} f(-y)$ ou bien existent les deux, ou bien n'existent pas, et si elles existent, elles se confondent:

$$\lim_{x \rightarrow \infty} f(x) = \lim_{y \rightarrow -\infty} f(-y).$$

d. Choisissons pour E l'intervalle $|x - x_0| \leq 1$ dépourvu du point x_0 , et pour F l'ensemble $|y| \geq 1$. La formule $y = \frac{1}{x - x_0}$ définit une application bijective de E sur F . A la direction $x \rightarrow x_0$ sur l'intervalle E correspond la direction $|y| \rightarrow \infty$ sur l'ensemble F . D'après a, les limites $\lim_{|y| \rightarrow \infty} f(y)$ et $\lim_{x \rightarrow x_0} f\left(\frac{1}{x - x_0}\right)$ existent ou n'existent pas les deux à la fois, et si elles existent, elles se confondent:

$$\lim_{|y| \rightarrow \infty} f(y) = \lim_{x \rightarrow x_0} f\left(\frac{1}{x - x_0}\right).$$

4.18. La définition d'une limite donnée dans 4.12 dépend, naturellement, de la métrique de l'espace M . Tout de même, en remplaçant la métrique $\rho(x, y)$ par une métrique homéomorphe $r(x, y)$ (3.34.c), la relation

$$\lim_s f(x) = p \quad (1)$$

reste valable pour la métrique $r(x, y)$ si elle l'était pour la métrique $\rho(x, y)$. Appliquons le critère d'homéomorphisme de 3.34.d. Supposons que la relation (1) soit valable pour la métrique ρ et que la métrique r soit homéomorphe à la métrique ρ . Soit un $\varepsilon > 0$; nous avons à trouver un $A \in S$ dont chaque point vérifie l'inégalité

$$r(f(x), p) < \varepsilon. \quad (2)$$

En vertu du critère 3.34.d, nous pouvons trouver, à partir d'un $\varepsilon > 0$ donné, un $\delta > 0$ tel que $\rho(p, y) < \delta$ implique $r(p, y) < \varepsilon$. D'après δ obtenu, choisissons un $A \in S$ de façon à avoir

$$\rho(f(x), p) < \delta$$

pour $x \in A$. Alors (2) est satisfait aussi par $x \in A$, ce qu'il fallait démontrer. Il est évident que la réciproque est vraie elle aussi: si la relation (2) est juste pour la métrique r , la métrique ρ étant

homéomorphe à la métrique r , alors la relation (1) est juste pour la métrique ρ .

4.19. Critère de Cauchy. Le critère de Cauchy d'existence de limite d'une suite numérique donné dans 3.72.a peut être étendu au cas général dans l'hypothèse que la fonction $f(x)$ prenne ses valeurs dans un espace complet.

Théorème. *Pour qu'une fonction $f(x)$ définie sur un ensemble E avec une direction S et à valeurs dans un espace métrique complet M ait une limite suivant S , il faut et il suffit que le critère de Cauchy soit rempli: pour tout $\varepsilon > 0$, il existe un ensemble $B \in S$ tel que*

$$\rho[f(x'), f(x'')] < \varepsilon \quad (1)$$

dès que $x' \in B$, $x'' \in B$.

Démonstration. Supposons rempli le critère de Cauchy. Considérons, dans l'espace M , tous les ensembles $f(A)$, où $A \in S$. (L'ensemble $f(A)$ est celui des valeurs que la fonction $f(x)$ prend sur l'ensemble A .) Les ensembles $A \in S$ formant un système d'ensembles emboîtés, les ensembles $f(A)$ en constituent un autre (sur M). Comme on le voit de l'inégalité (1), il y a, parmi les ensembles $f(A)$, des ensembles d'un diamètre autant petit que l'on veut. Donc, l'espace M étant complet et compte tenu du lemme 3.74.a, il existe dans M un point p tel que toute boule $V_\varepsilon(p) = \{y \in M : \rho(y, p) < \varepsilon\}$ contient entièrement l'un des ensembles $f(B)$, $B \in S$, de sorte que, pour tout $x \in B$, on a

$$\rho(f(x), p) < \varepsilon.$$

Or, ceci veut dire que la fonction $f(x)$ possède une limite égale à p suivant S .

La réciproque se passe de l'hypothèse de complétude de l'espace M . Soit $f(x)$ une fonction ayant une limite p suivant S . Trouvons, pour un $\varepsilon > 0$ donné, un ensemble $B \in S$ dont tout point satisfait à l'inégalité

$$\rho[p, f(x)] \leq \frac{\varepsilon}{2};$$

alors, quels que soient x' et x'' de B , nous avons

$$\rho[f(x'), f(x'')] \leq \rho[p, f(x')] + \rho[p, f(x'')] \leq \varepsilon,$$

ce qu'il nous faut.

§ 4.2. Théorèmes généraux sur les limites

Soient toujours E un ensemble quelconque portant une direction S et $f(x)$ une fonction définie sur E à valeurs dans un espace métrique M .

4.21. Si la fonction $f(x)$ est constante sur E , c'est-à-dire qu'elle prenne une même valeur $p \in M$ en tout point de l'ensemble E , alors sa limite suivant S existe et vaut p .

Evidemment, l'inégalité 4.12 (1) est, dans ce cas, satisfaite pour tout $\varepsilon > 0$ et sur tout $A \in S$.

4.22. La fonction $f(x)$ ne peut avoir qu'une seule limite suivant S . En effet, soit

$$p = \lim_S f(x) \in M \quad \text{et} \quad q = \lim_S f(x) \in M.$$

Pour un $\varepsilon > 0$ donné, trouvons un ensemble $A \subset S$ dont tout point vérifie l'inégalité

$$\rho(f(x), p) \leq \varepsilon, \quad (1)$$

et un ensemble $B \subset S$ dont tout point vérifie l'inégalité

$$\rho(f(x), q) \leq \varepsilon. \quad (2)$$

Soit, par exemple, $A \subset B$. Alors, pour tout point de l'ensemble A , les deux inégalités (1) et (2) sont satisfaites. Soit $x \in A$ un point quelconque, alors il résulte de (1) et (2) :

$$\rho(p, q) \leq \rho(f(x), p) + \rho(f(x), q) \leq 2\varepsilon.$$

Donc, $\varepsilon > 0$ étant arbitrairement petit, nous avons $\rho(p, q) = 0$, c'est-à-dire $p = q$.

4.23. On dit que la fonction $f(x)$ appartient asymptotiquement à un ensemble $G \subset M$ s'il existe un ensemble $A \in S$ en tout point duquel $f(x)$ appartient à G .

Théorème. Si $\lim_S f(x) = p$, et un ensemble $G \subset M$ contient une boule ouverte centrée en p , alors la fonction $f(x)$ appartient asymptotiquement à G .

Démonstration. Supposons que G contienne une boule $V = \{y : \rho(y, p) < \varepsilon\}$. Trouvons un ensemble $A \in S$ dont tout point vérifie l'inégalité $\rho(f(x), p) < \varepsilon$. Ceci veut dire que l'inclusion $f(x) \in V \subset G$ est valable elle aussi en tout point de l'ensemble A , ce qu'il fallait démontrer.

§ 4.3. Limites de fonctions numériques

4.31.a. Dans les paragraphes 4.3-4.6, nous construirons une théorie des limites de fonctions numériques, i.e. de fonctions qui prennent leurs valeurs sur la droite numérique (achevée ou non). Les particularités de telles fonctions sont dues à l'existence, sur la

droite numérique, d'une métrique, d'un ordre et des opérations arithmétiques. A vrai dire, il s'agit même de deux métriques différentes : la métrique usuelle $p(x, y) = |x - y|$ valable dans le domaine R des nombres finis, et la métrique $r(x, y)$ (3.35.e) valable dans l'ensemble achevé $\bar{R}$. Or, dans le domaine R , ces métriques sont homéomorphes (3.35.e); il en résulte (4.18) que l'existence d'une limite finie

$$p = \lim_s f(x)$$

ne dépend pas du choix de la métrique.

b. Soient $f(x)$ et $g(x)$ deux fonctions définies sur un ensemble E à valeurs dans $\bar{R}$. Nous notons $f(x) \leq g(x)$ si l'inégalité $f(x_0) \leq g(x_0)$ est remplie pour tout $x_0 \in E$. La fonction $f(x)$ est dite *majorée* sur l'ensemble E s'il existe un nombre fini C tel que, pour tout $x \in E$, on a l'inégalité $f(x) \leq C$. La fonction $f(x)$ est dite *minorée* sur E s'il existe un nombre fini C tel que, pour tout $x \in E$, on a l'inégalité $f(x) \geq C$. La fonction $f(x)$ est dite *bornée* (ou *bornée en module*) s'il existe un nombre fini C tel que, pour tout $x \in E$, on a l'inégalité $|f(x)| \leq C$.

En tout point $x_0 \in E$ où les deux fonctions $f(x)$ et $g(x)$ sont finies, nous définissons la somme $f(x) + g(x)$ comme résultat de l'addition des valeurs correspondantes $f(x_0)$ et $g(x_0)$. Soustraction, multiplication et division sont définies de façon analogue; la dernière, pour un point donné x_0 , n'est possible que si le dénominateur ne s'annule pas en ce point.

4.32. En développant la définition 4.23 nous dirons qu'une fonction numérique $f(x)$ est

- (a) *non négative (positive) suivant une direction S ,*
- (b) *bornée (majorée, minorée) suivant S ,*
- (c) *infinitement petite suivant S ,*
- (d) *infinitement grande positive suivant S ,*
- (e) *infinitement grande négative suivant S , si, respectivement,*

(a) il existe un ensemble $A \in S$ sur lequel la fonction $f(x)$ est non négative (positive);

(b) il existe un ensemble $B \in S$ sur lequel la fonction $f(x)$ est bornée (majorée, minorée); on note $f(x) = O(1)$;

(c) $\lim_s f(x) = 0$ (c'est-à-dire que, pour tout $\varepsilon > 0$, il existe un ensemble $B \in S$ sur lequel $|f(x)| < \varepsilon$); on note $f(x) = o(1)$;

(d) pour tout $C \in R$, il existe un ensemble $B \in S$ sur lequel $f(x) > C$; on note $\lim_s f(x) = \infty$;

(e) pour tout $C \in R$, il existe un ensemble $B \in S$ sur lequel $f(x) < C$; on note $\lim_s f(x) = -\infty$.

4.33. La relation $\lim_S f(x) = p$ est équivalente à la relation

$$\lim_S |f(x) - p| = 0.$$

Ceci résulte des définitions d'une limite et de la distance sur la droite numérique R :

$$\rho(a, b) = |a - b|.$$

4.34. a. L e m m e. Soit $\lim_S f(x) = p$; pour tout $\varepsilon > 0$, il existe un $A \in S$ dont chaque point vérifie les inégalités:

$$|f(x) - p| < \varepsilon \quad \text{si } p \text{ est fini,}$$

$$f(x) > \frac{1}{\varepsilon} \quad \text{si } p = +\infty,$$

$$f(x) < -\frac{1}{\varepsilon} \quad \text{si } p = -\infty.$$

Le lemme est une conséquence immédiate du théorème 4.33, compte tenu de la métrique en question (sur R et sur $\bar{R}$).

b. Soit un $p \neq 0$ fini; on peut trouver un $\varepsilon > 0$ tel que l'intervalle $(p - \varepsilon, p + \varepsilon)$ ne contient pas le point 0 et se compose donc de nombres de même signe que p lui-même. D'où:

Si une fonction $f(x)$ possède une limite $p > 0$ suivant une direction S , alors $f(x)$ est positive suivant S ; si $f(x)$ possède une limite $p < 0$ suivant S , alors $f(x)$ est négative suivant S .

Le résultat est valable pour $p = \infty$ et $p = -\infty$, ce qui s'ensuit directement de 4.34.a.

e. La proposition suivante est une sorte de réciproque de la proposition b:

Si $f(x)$ est non négative suivant la direction S et a p pour limite, alors $p \geq 0$.

En effet, si l'on avait $p < 0$, alors, d'après la propriété a, il existerait un ensemble $A \in S$ sur lequel $f(x) < 0$. D'autre part, $f(x)$ étant non négative suivant S , il existe un ensemble $B \in S$ sur lequel $f(x) \geq 0$. Comme AB n'est pas vide, nous aboutissons à une contradiction.

d. Soulignons que le passage « de la limite » (b) conserve le signe > 0 (ou bien < 0), tandis que le passage « à la limite » (c) conserve le signe ≥ 0 (ou bien ≤ 0). Si l'on sait seulement que $p \geq 0$, on ne peut rien conclure sur le signe de la fonction $f(x)$ sur les ensembles $A \in S$: elle peut prendre des valeurs positives aussi bien que négatives. Si l'on sait que $f(x) > 0$ sur certains $A \in S$, la limite p peut être positive aussi bien que nulle.

e. Nous dirons que $f(x) \leq g(x)$ ($f(x) < g(x)$) suivant une direction S si la différence $g(x) - f(x)$ est non négative (positive) suivant S .

f. Si $f(x)$ et $g(x)$ ont p et q respectivement pour limites suivant une direction S , et que $f(x) \leq g(x)$ suivant S , alors $p \leq q$. (Ceci résulte de la définition et de c.)

g. Si $f(x)$ et $g(x)$ ont p et q pour limites et que $p < q$, alors $f(x) < g(x)$ suivant S . (Ceci résulte de b.)

h. Si $f(x) \leq g(x) \leq h(x)$ suivant S et si $f(x)$ et $h(x)$ possèdent une même limite $p = \lim_S f(x) = \lim_S h(x)$, alors $g(x)$ possède aussi la limite p suivant S .

Il suffit d'examiner le cas $p = 0$ en remplaçant, dans le cas contraire, $f(x)$ et $g(x)$ par $f(x) - p$ et $g(x) - p$ (4.33). Pour un $\varepsilon > 0$ donné, trouvons un ensemble $A \in S$ sur lequel $|f(x)| < \varepsilon$ et $|h(x)| < \varepsilon$; il est évident que, sur l'ensemble A , on a aussi $|g(x)| < \varepsilon$, ce que nous cherchions.

4.35.a. Si $f(x)$ et $g(x)$ sont bornées suivant une direction S , alors $f(x) + g(x)$ et $f(x) \cdot g(x)$ le sont aussi.

En effet, supposons que l'inégalité $|f(x)| \leq C_1$ soit satisfaite sur un ensemble $A \in S$, et l'inégalité $|g(x)| \leq C_2$ sur un ensemble $B \in S$. Si, par exemple, $A \subset B$, alors, pour tout $x \in A$, on a les inégalités

$$|f(x) + g(x)| \leq C_1 + C_2 \quad \text{et} \quad |f(x)g(x)| \leq C_1C_2,$$

ce qu'il fallait démontrer.

b. Si $f(x)$ et $g(x)$ sont infiniment petites suivant S , alors $f(x) + g(x)$ l'est aussi.

Soient $A \in S$ et $B \in S$ des ensembles trouvés d'après un $\varepsilon > 0$ donné, sur lesquels $|f(x)| < \frac{\varepsilon}{2}$, $|g(x)| < \frac{\varepsilon}{2}$ respectivement; si, par exemple, $A \subset B$, alors

$$|f(x) + g(x)| < \varepsilon,$$

pour tout $x \in A$, ce qu'il nous fallait.

c. Si $f(x)$ est bornée suivant S et $g(x)$ est infiniment petite suivant S , alors $f(x)g(x)$ est infiniment petite suivant S .

En effet, supposons que l'inégalité $|f(x)| \leq C$ soit satisfaite sur un ensemble $A \in S$ et que, pour un $\varepsilon > 0$ donné, on ait un ensemble $B \in S$ sur lequel $|g(x)| \leq \frac{\varepsilon}{C}$. Sur AB on a l'inégalité $|f(x)g(x)| \leq \varepsilon$, ce que nous cherchions.

4.36.a. Si $\lim_S f(x) = p$, $\lim_S g(x) = q$, alors la fonction $f(x) + g(x)$ a sur S une limite qui vaut $p + q$.

En effet, $f(x) + g(x) - (p + q) = (f(x) - p) + (g(x) - q)$, la somme d'infiniment petits, est infiniment petite (4.35.b).

b. Si $\lim_S f(x) = p$, $\lim_S g(x) = q$, alors la fonction $f(x)g(x)$ a sur S une limite qui vaut pq .

En effet, $f(x)g(x) - pq = fg - fq + fq - pq = f(g - q) + q(f - p)$; chaque terme est un infiniment petit d'après 4.35.c.

e. Si $\lim_S f(x) = p \neq 0$, alors $\frac{1}{f(x)}$ est bornée suivant S .

En effet, trouvons un ensemble $A \in S$ sur lequel $|f(x)| \geq \frac{p}{2}$ (4.34.a); pour $x \in A$ nous avons $\frac{1}{|f(x)|} \leq \frac{2}{p}$, ce qu'il fallait démontrer.

d. Si $\lim_S f(x) = p \neq 0$, alors $\frac{1}{f(x)}$ a une limite valant $\frac{1}{p}$ suivant S .

En effet, $\frac{1}{p} - \frac{1}{f} = \frac{f-p}{pf} = \frac{1}{pf}(f-p)$ est un infiniment petit d'après c, 4.33 et 4.35.c.

e. Si $\lim_S g(x) = q$, $\lim_S f(x) = p \neq 0$, alors $\frac{g(x)}{f(x)}$ a la limite $\frac{q}{p}$ suivant S .

C'est une conséquence de b et d.

4.37.a. Si $\lim_S f(x) = 0$, alors $\lim_S \frac{1}{|f(x)|} = \infty$ et, réciproquement: si $\lim_S |g(x)| = \infty$, alors $\lim_S \frac{1}{g(x)} = 0$.

La conclusion découle de l'équivalence des inégalités $|f(x)| < \varepsilon$ et $\frac{1}{|f(x)|} > \frac{1}{\varepsilon}$.

b. Si $\lim_S f(x) = \infty$, alors $\lim_S [-f(x)] = -\infty$. Si $\lim_S f(x) = \infty$ et $\lim_S g(x) = p > 0$, alors $\lim_S f(x)g(x) = \infty$; si $\lim_S g(x) = p < 0$, alors $\lim_S f(x)g(x) = -\infty$. Ce sont des conséquences immédiates des définitions 4.32 (d), (e).

4.38. Définition. Soient $f(x)$ et $g(x)$ deux fonctions données sur un ensemble E avec une direction S . La fonction $f(x)$ est dite *infiniment petite par rapport à $g(x)$* (ou bien, si $g(x)$ est infiniment petite elle-même, *infiniment petite d'ordre supérieur par rapport à $g(x)$*) si

$$\lim_S \frac{f(x)}{g(x)} = 0.$$

On le note

$$f(x) = o(g(x)).$$

La fonction $f(x)$ est dite *infiniment grande par rapport à $g(x)$* (ou bien, si $g(x)$ est infiniment grande elle-même, *infiniment grande*

d'ordre supérieur par rapport à $g(x)$ si

$$\lim_s \left| \frac{f(x)}{g(x)} \right| = \infty.$$

Les fonctions $f(x)$ et $g(x)$ sont dites *équivalentes* (suivant S) si

$$\lim_s \frac{f(x)}{g(x)} = 1.$$

a. Il découle de 4.37.a que si $f(x)$ est infiniment petite par rapport à $g(x)$, alors $g(x)$ est infiniment grande par rapport à $f(x)$.

b. Si $f(x)$ et $g(x)$ sont équivalentes et s'il existe la limite

$$\lim_s \frac{g(x)}{h(x)} = L,$$

où $h(x)$ est une fonction sur E , alors la limite

$$\lim_s \frac{f(x)}{h(x)}$$

existe et vaut L elle aussi. En effet,

$$\frac{f(x)}{h(x)} = \frac{g(x)}{h(x)} \frac{f(x)}{g(x)} \rightarrow L$$

d'après 4.36.b. Donc, en cherchant la limite d'un rapport on peut remplacer son numérateur (et son dénominateur aussi) par une fonction équivalente.

4.39. S y m b o l e E .

a. Le symbole E désigne une fonction dont la limite (suivant une direction donnée) vaut 1. (Il serait plus correct d'écrire $E(x)$, mais on emploie une écriture abrégée.) Une fonction E , en accord avec 4.23, est appelée *unité asymptotique*. Evidemment, un produit et un quotient de deux unités asymptotiques sont des unités asymptotiques.

b. L e m m e. Si $\lim_s u(x) = 0$, alors, pour tout p naturel,

$$(1 + u(x))^p = 1 + pEu(x). \quad (1)$$

D é m o n s t r a t i o n. D'après la formule de Newton

$$\begin{aligned} (1 + u(x))^p &= 1 + pu(x) + \frac{p(p-1)}{1 \cdot 2} u^2(x) + \dots + u^p(x) = \\ &= 1 + pu(x) \left[1 + \frac{p-1}{1 \cdot 2} u(x) + \dots + \frac{1}{p} u^{p-1}(x) \right] = \\ &= 1 + pu(x) \cdot E, \end{aligned}$$

la limite de l'expression entre parenthèses valant 1 (4.36.a, b).

c. On se sert souvent d'unités asymptotiques en calculant des limites. Calculons, par exemple,

$$\lim_{x \rightarrow 0} \frac{(1+x)^p - (1+x+x^2)^q}{(1+2x)^r - (1-2x+x^3)^s} \quad (p, q, r, s \text{ des nombres naturels}).$$

D'après la définition et le lemme b, nous avons

$$\begin{aligned} (1+x)^p &= 1 + px E_1; \quad x+x^2 = x(1+x) = x E_2; \\ (1+x+x^2)^q &= (1+x E_2)^q = 1 + qx E_2 E_3 = 1 + qx E_4, \\ (1+2x)^r &= 1 + 2xr E_5, \quad 2x-x^3 = 2x \left(1 - \frac{1}{2} x^2\right) = 2x E_6; \\ (1-2x+x^3)^s &= (1-2x E_6)^s = 1 - 2xs E_6 E_7 = 1 - 2xs E_8; \\ \frac{(1+x)^p - (1+x+x^2)^q}{(1+2x)^r - (1-2x+x^3)^s} &= \frac{(1+px E_1) - (1+qx E_4)}{(1+2xr E_5) - (1-2xs E_8)} = \\ &= \frac{p E_1 - q E_4}{2r E_5 + 2s E_8} \rightarrow \frac{p-q}{2(r+s)}, \end{aligned}$$

ce qui achève le calcul.

d. Limite d'une fonction rationnelle $R(x)$ pour $x \rightarrow \pm\infty$. Considérons une fonction rationnelle

$$R(x) = \frac{a_0 x^m + a_1 x^{m-1} + \dots + a_m}{b_0 x^n + b_1 x^{n-1} + \dots + b_n}, \quad a_0 \neq 0, b_0 \neq 0.$$

Pour $|x| \rightarrow \infty$, nous avons

$$\begin{aligned} R(x) &= \frac{a_0 x^m \left(1 + \frac{a_1}{a_0} \frac{1}{x} + \dots + \frac{a_m}{a_0} \frac{1}{x^m}\right)}{b_0 x^n \left(1 + \frac{b_1}{b_0} \frac{1}{x} + \dots + \frac{b_n}{b_0} \frac{1}{x^n}\right)} = \\ &= \frac{a_0 x^m E_1}{b_0 x^n E_2} = \frac{a_0}{b_0} x^{m-n} E_2. \end{aligned}$$

Vu les règles 4.36-4.37, il en résulte pour $m \leq n$:

$$\lim_{|x| \rightarrow \infty} R(x) = \begin{cases} 0 & \text{pour } m < n, \\ \frac{a_0}{b_0} & \text{pour } m = n. \end{cases}$$

Lorsque $m > n$, la fonction $R(x)$ a une limite infinie dont le signe dépend de la direction choisie ($x \rightarrow +\infty$ ou $-\infty$), du signe de $\frac{a_0}{b_0}$ et de la parité ou de l'imparité du nombre $m - n$. Nous laissons au lecteur le soin de classer les cas possibles.

§ 4.4. Valeurs d'adhérence d'une fonction

4.41. Etudions le comportement d'une fonction numérique $f(x)$ à valeurs dans l'ensemble des nombres réels achevé $\bar{R}$ (§ 1.9) suivant une direction S . Pour tout ensemble $A \in S$, posons

$$a_A = \inf \{f(x) : x \in A\}, \quad b_A = \sup \{f(x) : x \in A\};$$

dans l'ensemble $\bar{R}$, les deux quantités indiquées existent, et $a_A \leq b_A$. Deux ensembles A et B quel onques du système S étant inclus l'un dans l'autre, il en est de même de n'importe quels deux intervalles $[a_A, b_A]$, $[a_B, b_B]$. Posons

$$\xi = \sup_{A \in S} a_A, \quad \eta = \inf_{A \in S} b_A; \quad (1)$$

d'après le principe généralisé de Cantor (1.94), $\xi \leq \eta$ et

$$[\xi, \eta] = \bigcap_{A \in S} [a_A, b_A]. \quad (2)$$

Le nombre $\xi \in \bar{R}$ s'appelle *valeur inférieure d'adhérence* (ou bien *limite inférieure*) de la fonction $f(x)$ suivant la direction S , le nombre $\eta \in \bar{R}$ *valeur supérieure d'adhérence* (ou bien *limite supérieure*) de la fonction $f(x)$ suivant S ; ceci se note :

$$\xi = \varliminf_S f(x), \quad \eta = \varlimsup_S f(x).$$

4.42. L'égalité (2) (4.41) entraîne le résultat suivant : pour tout $\varepsilon > 0$, il existe un ensemble $A \in S$ tel que

$$\begin{aligned} a_A &= \inf \{f(x) : x \in A\} \geq \xi - \varepsilon, \\ b_A &= \sup \{f(x) : x \in A\} \leq \eta + \varepsilon. \end{aligned}$$

(Si ξ appartient à $\bar{R} - R$, nous posons $\xi - \varepsilon = \xi$; d'une manière analogue, $\eta + \varepsilon = \eta$ si $\eta \in \bar{R} - R$.)

4.43. En conséquence, nous obtenons : si $\xi = \eta$, la fonction $f(x)$ a une limite suivant S qui vaut $\xi = \eta$; si ξ et η sont finis et $\xi \neq \eta$, alors $\xi = \eta$ est la limite finie de la fonction $f(x)$ suivant S .

4.44. Réciproquement, si la fonction $f(x)$ a la limite p suivant S , alors l'intervalle $[\xi, \eta] = \bigcap_{A \in S} [a_A, b_A]$ se réduit au point p , de sorte que $\xi = \eta = p$.

4.45. Pour élucider le sens des notions introduites, donnons encore la définition suivante (cf. la définition 3.41) :

Un nombre $y \in \bar{R}$ s'appelle *valeur d'adhérence de la fonction $f(x)$ suivant S* si, pour tout $\varepsilon > 0$ et pour tout $A \in S$, il existe un point $x \in A$ tel que

$$|f(x) - y| < \varepsilon \quad \text{si } y \text{ est fini;}$$

$$f(x) > \frac{1}{\varepsilon} \quad \text{si } y = \infty,$$

$$f(x) < -\frac{1}{\varepsilon} \quad \text{si } y = -\infty.$$

Lorsque la fonction $f(x)$ a une limite p suivant S , alors p est une valeur d'adhérence de la fonction $f(x)$ suivant S , mais la réciproque est, en général, fautive.

4.46. Montrons que, pour toute valeur d'adhérence $y \in R$ de la fonction $f(x)$ on a les inégalités $\xi \leq y \leq \eta$, ξ et η étant définis dans 4.41.

En effet, si y est une valeur d'adhérence de la fonction $f(x)$ suivant S , alors, pour tout $A \in S$:

$$a_A = \inf \{f(x) : x \in A\} \leq y \leq \sup \{f(x) : x \in A\} = b_A$$

donc, $y \in \bigcap [a_A, b_A] = [\xi, \eta]$.

Or, les points ξ et η eux-mêmes sont bien des valeurs d'adhérence de la fonction $f(x)$ suivant S , de sorte que la définition 4.41 ne contredit pas 4.45. Prouvons cela pour le point ξ . Si ξ est un nombre fini, alors, pour un $\varepsilon > 0$ donné, il existe un $A_0 \in S$ tel que $0 \leq$

$$\xi - a_{A_0} < \frac{\varepsilon}{2}, \text{ donc, pour tout } A \subset A_0, \text{ nous avons } a_{A_0} \leq$$

$$\leq a_A \leq \xi, \text{ d'où } 0 \leq \xi - a_A < \frac{\varepsilon}{2}. \text{ Puisque } a_A = \inf \{f(x) : x \in$$

$A\}$, il existe un point $x \in A$ pour lequel $0 \leq f(x) - a_A < \frac{\varepsilon}{2}$ et, par conséquent,

$$|f(x) - \xi| \leq \varepsilon.$$

Pour tout $B \supset A_0$, un tel point x existe aussi, et l'on peut en choisir un dans l'ensemble A_0 . Donc, ξ est une valeur d'adhérence de la fonction $f(x)$ suivant S . La démonstration pour le cas où $\xi \in \bar{R} - R$ ainsi que la démonstration analogue pour le point η suivent le même schéma.

4.47. C o n s é q u e n c e. L'ensemble de toutes les valeurs d'adhérence de la fonction $f(x)$ sur la droite achevée n'est pas vide et ses bornes sont les points $\xi = \liminf_S f(x)$ et $\eta = \limsup_S f(x)$.

4.48. Si la fonction $f(x)$ a une limite p suivant la direction S , alors, comme nous l'avons vu, $\xi = \eta = p$; dans ce cas, et seulement dans ce cas, le nombre p est l'unique valeur d'adhérence de la fonction $f(x)$ suivant la direction S .

§ 4.5. Fonctions non décroissantes suivant une direction

Une suite numérique $y_1, y_2, \dots, y_n, \dots$ est dite *non décroissante* si $y_1 \leq y_2 \leq \dots \leq y_n \leq y_{n+1} \leq \dots$.

Formulons une définition analogue pour une fonction numérique $y = f(x)$ donnée sur un ensemble quelconque E portant une direction S .

4.51. **Définition.** Une fonction numérique $f(x)$ est dite *non décroissante suivant une direction S* si, quels que soient les ensembles $A \in S$, $B \in S$, la relation $B \supset A$ entraîne l'inégalité

$$\sup \{f(x) : x \in B - A\} \leq \inf \{f(x) : x \in A\}. \quad (1)$$

4.52. Exemples.

a. Si $E = \{1, 2, \dots\}$, alors la fonction $y = f(x)$ est une suite numérique $y_1, y_2, \dots$. Soit, comme d'habitude, la direction S donnée sur E par les ensembles $A_n = \{n, n+1, \dots\}$, $n = 1, 2, \dots$. Montrons que la définition 4.51 se confond alors avec notre définition initiale d'une suite non décroissante.

Supposons que la suite f_n soit non décroissante au sens de la définition 4.51. Pour deux ensembles $A_n = \{n+1, \dots\}$ et $B_n = \{n, n+1, \dots\}$, nous avons $B_n - A_n = \{n\}$, de sorte que la condition 4.51 (1) implique

$$f_n \leq \inf \{f_p : p > n\} \leq f_{n+1}.$$

Réciproquement, il suit des inégalités $f_1 \leq f_2 \leq \dots \leq f_n \leq \dots$, quels que soient $A = \{f_n, f_{n+1}, \dots\} \in S$ et $A \subset B = \{f_k, f_{k+1}, \dots\} \in S$, $k < n$, que $\sup \{f(x) : x \in B - A\} = \sup \{f_k, \dots, f_{n-1}\} = f_{n-1} \leq f_n = \inf \{f(x) : x \in A\}$, de sorte que la suite $f(x)$ est non décroissante au sens de la définition 4.51.

b. Si $E = R_a^+$ est un demi-axe réel $a \leq x$ avec la direction S définie par les sous-ensembles $A_\xi = \{x \in R_a^+ : x \geq \xi\}$, alors la fonction $f(x)$ est non décroissante pour $x \rightarrow \infty$ si, et seulement si, $a \leq y < z$ implique $f(y) \leq f(z)$.

En effet, supposons que la fonction $f(x)$ soit non décroissante au sens de la définition 4.51. Alors, pour deux ensembles A_y et A_z , $A_y \supset A_z$, c'est-à-dire pour $y < z$, nous avons

$$f(y) \leq \sup \{f(x) : y \leq x < z\} \leq \inf \{f(x) : x \geq z\} \leq f(z)$$

Réciproquement, si l'on a $f(y) \leq f(z)$ quels que soient $y < z$ de R_1 , la fonction $f(x)$ est non décroissante au sens de la définition 4.51, puisque, si $x' \in A_y - A_z$ (c'est-à-dire $y \leq x' < z$) et $x'' \in A_z$ (c'est-à-dire $x'' \geq z$), alors $f(x') \leq f(x'')$ et, d'après 1.62.b,

$$\sup \{f(x') : x' \in (A_y - A_z)\} \leq \inf \{f(x'') : x'' \in A_z\}.$$

4.53. Théorème. Si une fonction $f(x)$ est non décroissante et majorée suivant une direction S , alors elle possède une limite suivant S : $\lim_S f(x) = \sup \{f(x) : x \in A\}$, où $A \in S$ est un ensemble quelconque sur lequel $f(x)$ est majorée.

Démonstration. Supposons que $f(x)$ soit majorée sur un ensemble $A \in S$. Posons $p = \sup \{f(x) : x \in A\}$. Pour un $\varepsilon > 0$ donné, trouvons un point $x_0 \in A$ tel que

$$p - \varepsilon < f(x_0) \leq p.$$

L'intersection de tous les ensembles du système S étant vide, il existe un ensemble $B \in S$ qui ne contient pas le point x_0 . Il est alors évident que c'est la première des deux inclusions possibles $A \supset B$, $B \supset A$ qui se réalise. Comme $f(x)$ ne décroît pas suivant S , on a

$$\inf \{f(x) : x \in B\} \geq \sup \{f(x) : x \in A - B\} \geq f(x_0) > p - \varepsilon.$$

Or, il résulte de $B \subset A$ que

$$\sup \{f(x) : x \in B\} \leq \sup \{f(x) : x \in A\} = p$$

(cf. 1.62.a). Donc,

$$p - \varepsilon < f(x) \leq p$$

partout sur B , ce qui achève la démonstration.

Pour une fonction non décroissante $f(x)$ possédant une limite p suivant S , on utilise la notation $f(x) \nearrow_S p$ ou tout simplement $f(x) \nearrow p$.

4.54. Une fonction numérique $f(x)$ est dite *non croissante suivant une direction S* si quels que soient les ensembles $A \in S$, $B \in S$, la relation $B \supset A$ entraîne l'inégalité

$$\inf \{f(x) : x \in B - A\} \geq \sup \{f(x) : x \in A\}.$$

Pour les fonctions non croissantes, un théorème analogue à 4.53 est valable :

Théorème. Si une fonction $f(x)$ est non croissante et minorée suivant une direction S , elle possède une limite suivant S :

$$p = \lim_S f(x) = \inf \{f(x) : x \in A\},$$

où $A \in S$ est un ensemble quelconque sur lequel $f(x)$ est minorée.

Ce théorème résulte immédiatement du théorème 4.53, dès que l'on remarque que, dans l'hypothèse de notre théorème, $-f(x)$ est non décroissante et majorée suivant S et que, si $p = \inf \{f(x) : x \in A\}$, alors, d'après 1.61,

$$\sup \{-f(x) : x \in A\} = -p.$$

Pour une fonction non croissante $f(x)$ qui possède une limite p , on utilise la notation $f(x) \searrow p$, ou tout simplement $f(x) \searrow p$.

4.55. Théorème. *Si une fonction $f(x)$ est non décroissante et non majorée suivant une direction S , alors*

$$\lim_S f(x) = \infty$$

(cf. 4.32 (d)).

En effet, soit un nombre C ; la fonction $f(x)$ n'étant pas majorée suivant la direction S , il existe, dans tout ensemble $B \in S$, un point $x_0 \in B$ tel que $f(x_0) > C$. Une fois B fixé, trouvons, dans la direction S , un ensemble $A \subset B$ ne contenant pas le point x_0 . Par hypothèse :

$$C \leq \sup \{f(x) : x \in B - A\} \leq \inf \{f(x) : x \in A\};$$

donc, l'inégalité

$$f(x) \geq C$$

est satisfaite en tout point de l'ensemble A , et notre théorème est démontré.

4.56. Un fait analogue a lieu pour une fonction $f(x)$ non croissante suivant une direction S :

Théorème. *Si une fonction numérique $f(x)$, est non croissante et non minorée suivant une direction S , alors $\lim_S f(x) = -\infty$.*

La démonstration consiste à appliquer 4.55 à la fonction $-f(x)$.

§ 4.6. Théorèmes fondamentaux sur des suites numériques

4.61. Nous allons appliquer la théorie générale des paragraphes 4.3-4.5 au cas de suites numériques. La définition d'une limite d'une suite numérique a déjà été donnée dans 3.32 (a). Rappelons-la. Une suite numérique $\{y_n\} = \{y_1, \dots, y_n, \dots\}$ ($y_n \in R$) est dite *convergente* vers un nombre $p \in R$ si, pour tout $\varepsilon > 0$, il existe un N naturel tel que l'inégalité

$$|p - y_n| < \varepsilon$$

est satisfaite dès que $n \geq N$.

Le nombre p est alors appelé *limite de la suite* y_n et se note :

$$p = \lim_{n \rightarrow \infty} y_n.$$

Lorsque les nombres y_n sont considérés comme points de l'espace achevé $\bar{R}$, la définition d'une limite finie reste la même, mais des limites infinies deviennent aussi possibles. Une suite $y_1, \dots, y_n, \dots$ qui converge vers $+\infty$ dans l'espace $\bar{R}$ est divergente dans l'espace R ; contrairement à d'autres suites divergentes, une telle suite est dite *divergente vers* $+\infty$. Une suite divergente vers $-\infty$ est définie d'une façon analogue.

4.62. Le critère de Cauchy 4.19 pour les suites numériques est confondu avec celui de 3.72:

Une suite de nombres $x_1, \dots, x_n, \dots$ possède une limite (finie) si, et seulement si, pour tout $\varepsilon > 0$, il existe un nombre N tel que

$$|x_m - x_n| < \varepsilon$$

quels que soient $m \geq N, n \geq N$.

4.63.a. Conformément à la définition 4.32 (b), une suite $x_1, \dots, x_n, \dots$ est dite *majorée* pour $n \rightarrow \infty$ (*minorée, bornée*) s'il existe un nombre C et un nombre N tels que, pour tout $n \geq N$, on a l'inégalité $x_n \leq C$ ($x_n \geq C, |x_n| \leq C$). D'ailleurs, ce n'est qu'un nombre fini de numéros de 1 à $N-1$ qui ne satisfont pas à la condition $n \geq N$; l'ensemble des nombres $x_1, \dots, x_{N-1}$ étant borné, on peut ne pas citer le nombre N dans la définition d'une suite bornée; on peut aussi bien omettre l'expression « pour $n \rightarrow \infty$ ».

b. Ensuite, conformément aux définitions 4.51 et 4.54, une suite $x_1, x_2, \dots$ est dite *non décroissante* (*non croissante*) pour $n \rightarrow \infty$ s'il existe un N tel que, pour $n \geq N$, l'inégalité $x_{n+1} \geq x_n$ ($x_{n+1} \leq x_n$) est satisfaite.

c. En vertu du théorème 4.53, toute suite $x_1, x_2, \dots$ majorée et non décroissante possède une limite pour $n \rightarrow \infty$. D'une façon analogue (4.54), une suite $x_1, x_2, \dots$ minorée et non croissante possède aussi une limite. Et si une suite non décroissante (non croissante) n'est pas majorée (minorée), elle converge vers ∞ ($-\infty$) (4.55, 4.56).

d. Ex e m p l e. Soit la suite numérique

$$u_n = \left(1 + \frac{1}{n}\right)^n \quad (n=1, 2, \dots).$$

Montrons que cette suite possède une limite finie pour $n \rightarrow \infty$. En effet, d'après la formule du binôme de Newton:

$$\begin{aligned} u_n &= \left(1 + \frac{1}{n}\right)^n = 1 + 1 + \frac{1}{2!} \frac{n(n-1)}{n^2} + \frac{1}{3!} \frac{n(n-1)(n-2)}{n^3} + \dots = \\ &= 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \frac{1}{3!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) + \dots; \end{aligned}$$

le nombre de termes de cette somme et chaque terme croissent avec n , donc les nombres u_n croissent eux aussi. Ensuite, en remplaçant $1 - \frac{k}{n}$ par 1, nous obtenons:

$$\begin{aligned} u_n &< 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} < \\ &< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} = 1 + \frac{1 - \frac{1}{2^n}}{1 - \frac{1}{2}} < 1 + 2 = 3, \end{aligned}$$

de sorte que la suite u_n est majorée. En appliquant c, nous voyons que la limite de la suite u_n existe. Cette limite est désignée par la lettre e :

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Ceci dit, il est clair que $2 < e \leq 3$. Un calcul plus exact montre que $e = 2,71828 \dots$. On peut démontrer que le nombre e est irrationnel (cf. l'exercice 14 du chapitre 8) et même transcendant (théorème de Hermite *)).

4.64. Citons d'autres définitions et faits concernant des valeurs d'adhérence des suites numériques. Conformément à la définition 4.45, un point $\xi \in \overline{R}$ s'appelle *valeur d'adhérence d'une suite* x_n si, pour tout $\varepsilon > 0$ et tout N , il existe un numéro $n \geq N$ tel que

$$\begin{aligned} |x_n - \xi| &< \varepsilon && \text{si } \xi \text{ est fini,} \\ x_n &> 1/\varepsilon && \text{si } \xi = \infty, \\ x_n &< -1/\varepsilon && \text{si } \xi = -\infty. \end{aligned}$$

Pour tout n , posons

$$\begin{aligned} a_n &= \inf \{x_n, x_{n+1}, \dots\} \in \overline{R}, \\ b_n &= \sup \{x_n, x_{n+1}, \dots\} \in \overline{R}. \end{aligned}$$

D'après 4.41, les intervalles $[a_n, b_n]$ de la droite numérique achevée forment une suite d'intervalles emboîtés avec $[\xi, \eta]$ pour intersection, où $\xi = \sup a_n$, $\eta = \inf b_n$. Toutes les valeurs d'adhérence de la suite x_n appartiennent à l'intervalle $[\xi, \eta]$; de plus, les points ξ et η eux-mêmes sont toujours des valeurs d'adhérence et s'appellent *limites inférieure et supérieure* respectivement de la suite; on le désigne comme suit:

$$\xi = \lim_{n \rightarrow \infty} a_n, \quad \eta = \lim_{n \rightarrow \infty} b_n.$$

*) Voir, par exemple, [8].

Lorsque $\xi = \eta$, la suite x_n a pour limite $\xi = \eta$ ($\in \bar{R}$). Lorsque $\xi = \eta$ tout en étant fini, la suite x_n a une limite finie $\xi = \eta$. Réciproquement, si la suite x_n a une limite p , alors $\xi = \eta = p$ et la suite x_n a le nombre p pour son unique valeur d'adhérence.

4.65. Limite d'une suite et limite d'une fonction. Si une fonction $f(t)$ définie pour $t \geq t_0$ possède une limite pour $t \rightarrow \infty$, disons $\lim_{t \rightarrow \infty} f(t) = p$, alors la suite des nombres

$$y_n = f(n) \quad (n \text{ entiers supérieurs à } t_0)$$

possède la même limite p (4.16.a).

La conclusion réciproque sur l'existence de la limite d'une fonction $f(x)$ définie pour $x \geq x_0$ à partir de l'existence de la limite de la suite $f(n)$ n'est pas valable. Ainsi, la fonction $f(x) = (x)$ (partie fractionnaire de x , 1.71) n'a pas de limite pour $x \rightarrow \infty$, bien que $f(n) = (n) = 0$ ait 0 pour limite.

4.66. Tout de même, le théorème suivant a lieu :

Théorème. Pour qu'une fonction $f(x)$ définie pour $x \geq x_0$ ait une limite, lorsque $x \rightarrow \infty$, il faut et il suffit que toute suite $f(x_n)$ (où $x_n \geq x_0$ est une suite quelconque divergente vers ∞) possède une limite.

Démonstration. 1) Si $f(x)$ a une limite p pour $x \rightarrow \infty$ alors toute suite $f(x_n)$, où $x_n \rightarrow \infty$, a la même limite p .

Ceci résulte toujours de 4.16.a.

2) Si $f(x)$ n'a pas de limite pour $x \rightarrow \infty$, alors le critère de Cauchy 4.19 n'est pas satisfait. Donc, pour un $h > 0$ et pour tout n , on peut trouver des points $x'_n > n$, $x''_n > n$ tels que

$$|f(x'_n) - f(x''_n)| > h. \quad (1)$$

Considérons la suite

$$x'_1, x''_1, x'_2, x''_2, \dots, x'_n, x''_n, \dots$$

Il est évident qu'elle tend vers ∞ et que la suite des valeurs correspondantes de la fonction $f(x)$:

$$f(x'_1), f(x''_1), f(x'_2), f(x''_2), \dots$$

n'a pas de limite puisque (1) montre que cette suite ne satisfait pas au critère de Cauchy.

Les faits exposés n'écartent pas la possibilité de conclure, dans des cas spéciaux, sur l'existence de la limite, pour $x \rightarrow \infty$, d'une fonction $f(x)$ à partir de l'existence de $\lim_{n \rightarrow \infty} f(n)$ si l'on utilise des propriétés supplémentaires de la fonction $f(x)$.

§ 4.7. Limites de fonctions vectorielles

4.71. Considérons ici les fonctions $f(x)$ à valeurs dans l'espace n -dimensionnel R_n (2.61). L'espace n -dimensionnel étant métrique, la définition d'une limite suivant une direction S est applicable, et, pour une telle limite, les propriétés données dans § 4.2 sont valables. De plus, certaines propriétés des limites de fonctions numériques mentionnées dans §§ 4.3-4.5 restent vraies pour des fonctions à valeurs dans R_n , notamment les propriétés liées aux opérations linéaires et n'utilisant pas l'ordre. Voyons-les brièvement.

4.72. Pour les fonctions à valeurs dans un espace n -dimensionnel R_n , les opérations suivantes sont définies :

a. **Addition** : si deux fonctions $f(x)$ et $g(x)$ prennent les valeurs dans l'espace R_n , leur somme, c'est-à-dire la fonction égale, pour tout $x_0 \in E$, à la somme $f(x_0) + g(x_0)$, est une fonction à valeurs dans R_n .

b. **Multiplication par une fonction réelle** : si une fonction $f(x)$ prend les valeurs dans l'espace R_n , et si une fonction $\alpha(x)$ a les valeurs réelles, leur produit, c'est-à-dire la fonction égale, pour tout $x_0 \in E$, au produit $\alpha(x_0) f(x_0)$, est une fonction à valeurs dans l'espace R_n .

c. En cas de $n = 2$, où les fonctions vectorielles $f(x)$ et $g(x)$ peuvent être interprétées comme fonctions à valeurs complexes,

le produit $f(x) \cdot g(x)$ et le quotient $\frac{f(x)}{g(x)}$ (le dernier pour $g(x) \neq 0$) sont définis d'après les règles d'opérations sur les nombres complexes.

4.73.a. La relation $\lim_S f(x) = p \in R_n$ est équivalente à $\lim_S |f(x) - p| = 0$ et à $\lim_S |f(x) - p| = 0$.

b. Une fonction $f(x)$ à valeurs dans R_n est dite *bornée suivant une direction S* s'il existe un nombre fini c et un ensemble $A \in S$ dont chaque point vérifie l'inégalité $|f(x)| \leq c$.

Si $f(x)$ et $g(x)$ sont bornées suivant une direction S , la fonction $f(x) + g(x)$ l'est aussi.

c. Si $\lim_S f(x) = p \in R_n$, $\lim_S g(x) = q \in R_n$, alors la fonction $f(x) + g(x)$ possède, suivant la direction S , une limite valant $p + q$:

$$\lim_S |f(x) + g(x)| = \lim_S f(x) + \lim_S g(x).$$

d. Si $\lim_S f(x) = p \in R_n$, $\lim_S \alpha(x) = l \in R$, alors la fonction $\alpha(x) \cdot f(x)$ possède, suivant la direction S , une limite qui vaut lp :

$$\lim_S \alpha(x) f(x) = \lim_S \alpha(x) \cdot \lim_S f(x).$$

e. En cas de $n = 2$, où les fonctions vectorielles $f(x)$, $g(x)$ peuvent être interprétées comme fonctions à valeurs complexes, de sorte que l'opération de multiplication a un sens, nous avons: si $\lim_S f(x) = p \in C$, $\lim_S g(x) = q \in C$, alors $\lim_S f(x) \cdot g(x) = pq$ et

$$\lim_S \frac{f(x)}{g(x)} = \frac{p}{q} \text{ (la dernière égalité pour } q \neq 0 \text{)}.$$

f. Exemple. Dans le plan C des nombres complexes $z = x + iy$ définissons la direction $z \rightarrow \infty$ formée de tous les ensembles A_r de la forme

$$A_r = \{z \in C : |z| > r\}.$$

Les propriétés d'une direction (4.11) sont, évidemment, remplies, et on peut discuter l'existence de la limite $\lim_{z \rightarrow \infty} f(z)$ pour des fonctions

$f(z)$ définies pour $|z| \geq r_0$. En particulier, si $f(z) = \frac{1}{z}$ pour $z \neq 0$, nous avons:

$$\lim_{z \rightarrow \infty} \frac{1}{z} = 0,$$

puisque, quel que soit $\varepsilon > 0$, nous avons $|f(z)| = \frac{1}{|z|} < \varepsilon$ sur l'ensemble

$$A_{1/\varepsilon} = \left\{z \in C : |z| > \frac{1}{\varepsilon}\right\}.$$

Ensuite, pour tout polynôme de la variable $\frac{1}{z}$, c'est-à-dire pour tout polynôme de la forme

$$a_0 + \frac{a_1}{z} + \dots + \frac{a_n}{z^n}$$

aux coefficients complexes $a_0, a_1, \dots, a_n$, on a, en vertu de c et e:

$$\lim_{z \rightarrow \infty} \left(a_0 + \frac{a_1}{z} + \dots + \frac{a_n}{z^n}\right) = a_0.$$

4.74. Une fonction $f(x)$ à valeurs dans un espace n -dimensionnel R_n peut être mise sous la forme

$$f(x) = (f_1(x), \dots, f_n(x)),$$

où $f_k(x)$ sont les coordonnées du vecteur $f(x)$ qui représentent des fonctions numériques. De plus, l'inégalité suivante a lieu (2.68 (7)) :

$$\begin{aligned} \max_{k=1, \dots, n} |f_k(x) - f_k(y)| &\leq \sqrt{\sum_{j=1}^n |f_j(x) - f_j(y)|^2} \leq \\ &\leq \sum_{j=1}^n |f_j(x) - f_j(y)|. \end{aligned} \quad (1)$$

Nous en tirons une conclusion importante :

Théorème. Une fonction $f(x) \in R_n$ a une limite suivant une direction S si, et seulement si, chacune des fonctions numériques $f_1(x), \dots, f_n(x)$ a une limite suivant la direction S .

4.75. Formulons, pour l'espace R_n , le critère de Cauchy (4.19) qui est valable, car R_n est complet (3.72.c).

Théorème (critère de Cauchy pour R_n). Pour qu'une fonction $f(x)$ à valeurs dans R_n ait une limite suivant une direction S , il faut et il suffit que, pour tout $\varepsilon > 0$, il existe un ensemble $A \in S$ tel que, quels que soient deux points x et y de A , on ait

$$|f(x) - f(y)| \leq \varepsilon. \quad (1)$$

Exercices

1. Démontrer que si une suite $x_n \in R$ est convergente, alors la suite $|x_n|$ l'est aussi. Est-ce que la réciproque est vraie?

2. Pour deux suites réelles quelconques a_n et b_n

$$\lim a_n + \overline{\lim} b_n \leq \overline{\lim} (a_n + b_n) \leq \overline{\lim} a_n + \overline{\lim} b_n,$$

$$\lim a_n + \overline{\lim} b_n \geq \lim (a_n + b_n) \geq \lim a_n + \lim b_n.$$

3. Si une suite a_n converge, alors toute sa permutation $a_{n_1}, a_{n_2}, \dots, a_{n_k}, \dots$ converge vers la même limite. Est-ce que la convergence d'une suite résulte de la convergence de l'une de ses permutations?

4. Si une suite a_n converge, alors, pour toute suite b_n , nous avons :

$$\overline{\lim} (a_n + b_n) = \overline{\lim} a_n + \overline{\lim} b_n.$$

5. Si, pour une suite a_n et pour toute suite b_n , on a

$$\overline{\lim} (a_n + b_n) = \overline{\lim} a_n + \overline{\lim} b_n,$$

alors la suite a_n converge.

6. Soit $x_1 = a$, $x_2 = b$, $x_3 = \frac{1}{2}(a+b)$, $\dots$, $x_n = \frac{1}{2}(x_{n-1} + x_{n-2})$, $\dots$. Trouver $\lim x_n$.

7. Soit $a > 0$, $x_0 > 0$ et

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right) \quad (n=0, 1, 2, \dots).$$

Démontrer que $\lim x_n = \sqrt{a}$.

8. (Moyenne arithmético-géométrique de Gauss). Soit $a > 0$, $b > 0$; posons $x_1 = a$, $y_1 = b$, ..., $x_{n+1} = \sqrt{x_n y_n}$, $y_{n+1} = \frac{x_n + y_n}{2}$, ...; démontrer que les nombres x_n et y_n ont une limite commune.

9. Si $\max\{p_1, \dots, p_m\} = p_1$ ($p_1 > 0, \dots, p_m > 0$), alors

$$\lim_{n \rightarrow \infty} \sqrt[n]{\sum_{k=1}^m p_k^n} = p_1.$$

10. Une droite $y = kx + b$ est asymptote à une courbe $y = f(x)$, pour $x \rightarrow +\infty$, si

$$\lim_{x \rightarrow +\infty} [f(x) - (kx + b)] = 0.$$

Démontrer que la courbe $y = f(x)$ possède une asymptote pour $x \rightarrow +\infty$ si les limites

$$\lim_{x \rightarrow \infty} \frac{f(x)}{x}, \quad \lim_{x \rightarrow +\infty} \left[f(x) - x \lim_{x \rightarrow +\infty} \frac{f(x)}{x} \right]$$

existent, et dans ce cas seulement.

11. Soit $f(x)$ une fonction définie sur un espace métrique M à valeurs dans un espace métrique P . Nous disons qu'un point $p \in P$ est une limite « renforcée » de la fonction $f(x)$ dans la direction $x \rightarrow a \in M$ si, quel que soit $\varepsilon > 0$, il existe un $\delta > 0$ tel que, pour n'importe quels $x \in M$ tels que $\rho(x, a) < \delta$ (sans exception le point $x = a$), l'inégalité $\rho(f(x), p) < \varepsilon$ est satisfaite. La limite « renforcée » n'entre pas directement dans le schéma 4.12 de la limite suivant une direction. Montrer que pour qu'un point $p \in P$ soit une limite « renforcée » de la fonction $f(x)$ pour $x \rightarrow a$, il faut et il suffit que p soit une limite de la fonction $f(x)$ suivant la direction $x \rightarrow a$ et que, de plus, $f(a) = p$.

12. Soit $y(x)$ une fonction définie sur un ensemble X portant une direction $S = \{A_\alpha\}$ et à valeurs dans un espace métrique Y ; soit $z(y)$ une fonction définie sur Y à valeurs dans un espace métrique Z . La fonction composée $z(x) = z[y(x)]$ est définie sur X et prend ses valeurs dans Z . Supposons que les limites

$$p = \lim_S y(x) \in Y, \quad q = \lim_{y \rightarrow q} z(y) \in Z$$

existent.

Peut-on dire que la limite $\lim_S z(x)$ existe elle aussi? Si oui, est-elle égale à q ?

13. Dans l'hypothèse de l'exercice 12, démontrer les propositions suivantes:

a) S'il existe un ensemble $A \in S$ sur lequel $y(x)$ ne prend pas la valeur p , alors $z(x)$ possède une limite q .

b) S'il existe un ensemble $A \in S$ sur lequel $y(x)$ est identiquement égale à p , alors $z(x)$ possède une limite $z(p)$.

c) Si, sur tout ensemble $A \in S$, la fonction $y(x)$ prend la valeur p aussi bien que des valeurs différentes de p , alors la fonction $z(x)$ a une limite si, et seulement si, $q = z(p)$ (c'est-à-dire que q est une limite « renforcée » de la fonction $z(y)$ pour $y \rightarrow p$) et dans ce cas:

$$\lim_S z(x) = q.$$

Histoire

La première définition correcte d'une limite d'une fonction numérique est donnée par Cauchy dans son « Cours d'analyse de l'Ecole polytechnique » (1821). Cauchy établit les théorèmes fondamentaux sur l'existence de limites de divers genres, en particulier celles de fonctions et suites monotones. C'est toujours lui qui introduit les notions de limites supérieure et inférieure. Les notions plus générales de limite sont proposées par Chatounovski (1923), Moore et Smith (1923). La convergence « suivant une direction » est un cas spécial de la convergence « suivant un filtre » de H. Cartan (1937) [46].

Fonctions continues

Mon but principal a été de concilier la rigueur, dont je m'étais fait une loi dans mon Cours d'analyse, avec la simplicité qui résulte de la considération directe des quantités infiniment petites

Au. Cauchy (1823)

§ 5.1. Fonctions continues sur un espace métrique

5.11. Soit $f(x)$ une fonction donnée sur un espace métrique M à valeurs dans un espace métrique P . Soit $a \in M$ un point fixe non isolé (4.15.a). Considérons la direction $x \rightarrow a$ (4.15.a) formée des boules $U_r(a) = \{x: \rho(x, a) < r\}$ dépourvues du point a .

a. Définition. La fonction $f(x)$ est dite *continue* pour $x = a$ (le point a est alors appelé *point de continuité* de la fonction $f(x)$) si

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Autrement dit, la fonction $f(x)$ est continue pour $x = a$ si, quel que soit $\varepsilon > 0$, il existe un $\delta > 0$ tel que $\rho(x, a) < \delta$ implique

$$\rho(f(x), f(a)) < \varepsilon. \quad (1)$$

Pour $f(x)$ numérique, l'inégalité (1) prend, naturellement, la forme :

$$|f(x) - f(a)| < \varepsilon.$$

b. Il résulte de 4.66 une autre définition de la continuité : la fonction $f(x)$ est continue pour $x = a$ si $f(x_n) \rightarrow f(a)$ (dans P) pour toute suite $x_n \rightarrow a$ (dans M).

c. Tout point isolé $a \in M$ est, par définition, un point de continuité de la fonction $f(x)$.

d. Définition. Tout point x_0 qui n'est pas un point de continuité de la fonction $f(x)$ s'appelle son *point de discontinuité*.

e. Définition. Une fonction continue en tout point de l'espace M est dite *continue sur M* .

f. La définition de la continuité de la fonction $f(x)$ en un point a dépend, naturellement, des métriques données sur les espaces M

et P . Cette définition pouvant être formulée dans les termes de suites convergentes (b), la propriété d'une fonction d'être continue en un point a , aussi bien que sur tout l'ensemble M , n'est pas violée en remplaçant les métriques des espaces M et P par les métriques homéomorphes (3.34).

5.12.a. Une fonction constante

$$f(x) = y_0,$$

y_0 étant un point fixe de l'espace P (4.21), sert d'exemple évident d'une fonction continue avec M pour domaine de définition et P pour domaine de valeurs.

b. En tant que deuxième exemple, considérons la distance $\rho(x, a)$ d'un point fixe a . C'est une fonction numérique sur l'espace métrique M . Sa continuité en tout point $x = x_0$ de l'espace M découle de 3.32.b.

5.13. Les propositions suivantes relatives aux fonctions numériques permettent de construire de vastes classes de fonctions continues;

a. Si deux fonctions numériques $f(x)$ et $g(x)$ sont continues pour $x = x_0$, alors $h(x) = f(x) + g(x)$ est aussi continue pour $x = x_0$. C'est une conséquence de 4.36.a.

b. Si $f(x)$ et $g(x)$ sont continues pour $x = x_0$, alors $p(x) = f(x) \cdot g(x)$ est aussi continue pour $x = x_0$. C'est une conséquence de 4.36.b.

c. Si $f(x)$ et $g(x)$ sont continues pour $x = x_0$ et $g(x_0) \neq 0$, alors $q(x) = \frac{f(x)}{g(x)}$ est aussi continue pour $x = x_0$.

C'est une conséquence de 4.36.c.

d. La fonction numérique $y = x$ définie sur la droite numérique R est, évidemment, continue sur R . Il résulte des propositions a, b et c que tout polynôme $a_0 + a_1x + \dots + a_nx^n$ est partout continu sur R et que toute fonction rationnelle $r(x) = \frac{a_0 + a_1x + \dots + a_nx^n}{b_0 + b_1x + \dots + b_mx^m}$ est partout continue sur R sauf en des points où son dénominateur s'annule.

e. La fonction numérique $y = \xi_k$ (k -ième coordonnée du vecteur $x = (\xi_1, \dots, \xi_n)$) est, évidemment, une fonction continue sur l'espace n -dimensionnel R_n . Il résulte des propositions a-c que tout polynôme par rapport aux coordonnées du vecteur x est partout continu sur R_n et toute fonction rationnelle des coordonnées du vecteur x est partout continue sur R_n , sauf en des points où son dénominateur s'annule.

5.14. Soit $f(x)$ une fonction continue définie sur un espace métrique M , à valeurs dans un espace métrique P . Soient, ensuite, $G \subset P$ un ensemble et $H = \{x \in M : f(x) \in G\}$.

a. Si G est ouvert, H l'est aussi.

En effet, soient $x_0 \in H$, $f(x_0) \in G$ et soit $V = \{y \in P : \rho(y, f(x_0)) < \varepsilon\}$ une boule de centre $f(x_0)$ dans le domaine G . Trouvons un $\delta > 0$ tel que $\rho(x, x_0) < \delta$ implique $\rho(f(x), f(x_0)) < \varepsilon$. Alors la boule $U = \{x : \rho(x, x_0) < \delta\}$ est dans H .

b. Si G est fermé, H l'est aussi.

En effet, l'ensemble $\{x : f(x) \in P - G\}$ qui est ouvert en vertu de a est le complémentaire de l'ensemble $\{x : f(x) \in G\}$ dans M .

c. Conséquence. Si $f(x)$ est une fonction numérique continue, alors, pour tout $c \in \bar{R}$ réel, les ensembles

$$\{x \in M : f(x) < c\},$$

$$\{x \in M : f(x) > c\}$$

sont ouverts, et les ensembles

$$\{x : f(x) \leq c\},$$

$$\{x : f(x) \geq c\},$$

$$\{x : f(x) = c\}$$

fermés.

d. Si $f_1(x), \dots, f_m(x)$ sont des fonctions numériques continues, alors, quels que soient $a_1, \dots, a_m, b_1, \dots, b_m$ réels (de $\bar{R}$), l'ensemble

$$\{x : a_1 < f_1(x) < b_1, \dots, a_m < f_m(x) < b_m\} \quad (1)$$

est ouvert, et l'ensemble

$$\{x : a_1 \leq f_1(x) \leq b_1, \dots, a_m \leq f_m(x) \leq b_m\} \quad (2)$$

fermé.

Ceci résulte de c compte tenu des théorèmes sur l'intersection d'ensembles ouverts (3.22) et fermés (3.54).

e. Etant donnée la proposition d, bien des figures de la géométrie élémentaire décrites par les systèmes d'inégalités (1) ou (2) s'avèrent respectivement ouvertes ou fermées. Ainsi, dans un espace n -dimensionnel, un ensemble de la forme

$$\{x : a_1 < x_1 < b_1, \dots, a_n < x_n < b_n\}$$

est ouvert, et un ensemble de la forme

$$\{x : a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\}$$

fermé; le premier est appelé *pavé ouvert*, le second *pavé fermé*.

Dans le plan x_1, x_2 , un m -gone décrit par m inégalités de la forme $a_j x_1 + b_j x_2 < c_j$ ($j = 1, \dots, m$) est un ensemble ouvert appelé *m-gone ouvert*; si l'on remplace, dans lesdites inégalités, le symbole $<$ par $\leq$, on obtient un *m-gone fermé*. Dans un espace n -dimensionnel, on définit d'une façon analogue les *polyèdres ouverts* et *fermés* à l'aide d'inégalités qui relient des fonctions linéaires des coordonnées.

Celles des figures mentionnées qui sont fermées et bornées en même temps sont des compacts en vertu de 3.96.e.

5.15. Continuité d'une fonction composée.

a. Soient M, N, P des espaces métriques avec pour distances ρ_M, ρ_N, ρ_P respectivement. Soient $y = f(x), z = \varphi(y)$ deux fonctions dont les domaines de définition et les domaines de valeurs sont respectivement M et N et N et P . La fonction composée $z = \varphi[f(x)] = h(x)$ est définie dans M et prend les valeurs dans P .

Théorème. *Si $y = f(x)$ est continue pour $x = x_0$ et $z = \varphi(y)$ l'est pour $y = y_0 = f(x_0)$, alors la fonction $z = h(x)$ est continue pour $x = x_0$.*

Démonstration. Soit un $\varepsilon > 0$. En partant de la condition $\lim_{y \rightarrow y_0} \varphi(y) = \varphi(y_0)$ trouvons un nombre $\tau > 0$ tel que $\rho_N(y, y_0) < \tau$ implique $\rho_P(\varphi(y), \varphi(y_0)) < \varepsilon$. Ensuite, à l'aide de la condition $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ et de τ trouvé, choisissons un $\delta > 0$ tel que $\rho_M(x, x_0) < \delta$ implique $\rho_N(y, y_0) = \rho_N(f(x), f(x_0)) < \tau$. Alors, pour $\rho_M(x, x_0) < \delta$, nous avons

$$\begin{aligned} \rho_P(h(x), h(x_0)) &= \rho_P(\varphi(f(x)), \varphi(f(x_0))) = \\ &= \rho_P(\varphi(y), \varphi(y_0)) < \varepsilon, \quad \text{c.q.f.d.} \end{aligned}$$

b. Si $f(x)$ est une fonction continue définie sur un espace métrique M , à valeurs dans un espace métrique N et si $b \in N$ est un point fixe, alors la fonction numérique

$$h(x) = \rho(f(x), b)$$

est une fonction continue de $x \in M$.

C'est une conséquence de a et de 5.12.b.

En particulier, si une fonction numérique $f(x)$ (à valeurs dans $\mathbb{R}$) est continue sur un espace métrique M , alors la fonction $|f(x)| = \rho(f(x), 0)$ est continue elle aussi.

5.16. Fonctions continues sur un compact. Les fonctions continues sur un compact métrique M (3.91.a) possèdent certaines propriétés spécifiques que nous étudions ici et dans 5.17.

a. **Théorème.** *L'ensemble de toutes les valeurs d'une fonction continue $f(x)$ définie sur un compact M est compact.*

Démonstration. Soit $f(x_1), \dots, f(x_n), \dots$ une suite quelconque de valeurs d'une fonction continue $f(x)$ définie sur un compact M . Extrayons de la suite des points $x_1, \dots, x_n, \dots$ une suite convergente x_{n_k} ($k = 1, 2, \dots$); par exemple, $x_{n_k} \rightarrow a$.

La fonction $f(x)$ étant continue pour $x = a$, ou $a f(x_n) \rightarrow f(a)$. Ainsi, on a extrait de la suite $f(x_n)$ une sous-suite convergente, ce qu'il fallait démontrer.

b. Conséquence (théorème de Weierstrass). Une fonction numérique continue (à valeurs dans R) définie sur un compact M est bornée et atteint sur M ses bornes, c'est-à-dire que si $\alpha = \inf f(x)$, $\beta = \sup f(x)$, alors il existe les points $q \in M$, $p \in M$ tels que $\alpha = f(q)$, $\beta = f(p)$.

La démonstration est immédiate si l'on part de 5.16.a et des propriétés d'un ensemble compact sur la droite numérique (3.96.f).

c. Dans l'hypothèse de b, choisissons pour compact M l'intervalle $a \leq x \leq b$ de la droite numérique, et soit $f(a) = f(b)$. On peut alors affirmer que l'un au moins des points p, q figurant dans le théorème de Weierstrass se trouve dans l'intervalle (a, b) . En effet, si la fonction $f(x)$ est constante et vaut $f(a)$, on peut choisir pour $q = p$ n'importe quel point de l'intervalle (a, b) . Supposons que $f(x)$ ne soit pas constante et que certaines de ses valeurs puissent être supérieures, par exemple, à $f(a)$. Alors $\beta = \sup f(x) > f(a)$. D'après le théorème de Weierstrass, il existe un point $p \in [a, b]$ pour lequel $f(p) = \beta$. Or $f(a) = f(b) < \beta$, donc $p \in (a, b)$. Si $f(x)$ prend des valeurs inférieures à $f(a)$, on fait le raisonnement analogue pour le point q en lequel $\inf f(x)$ est atteint.

5.17. Continuité uniforme.

a. Définition. Une fonction $y = f(x)$ définie sur un espace métrique M , à valeurs dans un espace métrique P est dite *uniformément continue* sur M si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $\rho(x_1, x_2) < \delta$ implique $\rho(f(x_1), f(x_2)) < \varepsilon$ quels que soient les points x_1 et x_2 de M .

Il est évident que toute fonction uniformément continue est continue en tout point de l'espace M .

Mais, dans le cas général, la continuité d'une fonction $f(x)$ sur un espace M n'implique pas sa continuité uniforme. Par exemple, la fonction $y = x^2$ est bien continue sur le demi-axe $0 \leq x < \infty$ (5.13.d) sans être uniformément continue sur lui parce que

$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2).$$

et la condition $|x_1 - x_2| < \delta$ n'implique sûrement pas que $|x_1^2 - x_2^2|$ est bornée par une constante.

b. Dans le cas où M est un compact métrique, le théorème suivant a lieu.

Théorème (Heine). Une fonction continue $y = f(x)$ définie sur un compact métrique K , à valeurs dans un espace métrique P est uniformément continue sur K .

Démonstration. Supposons que la fonction $f(x)$ ne soit pas uniformément continue sur K . Alors, pour un $\varepsilon > 0$ et pour tout $\delta > 0$, il existe deux points x', x'' dans K pour lesquels $\rho(x', x'') < \delta$, mais $\rho(f(x'), f(x'')) \geq \varepsilon$. Donnons au nombre δ les valeurs successives $\delta = 1, \frac{1}{2}, \dots, \frac{1}{n}, \dots$. Pour tout $\delta = \frac{1}{n}$, trouvons convenablement les points x'_n et x''_n de sorte que $\rho(x'_n, x''_n) < \delta = \frac{1}{n}$, mais

$$\rho[f(x'_n), f(x''_n)] \geq \varepsilon. \quad (1)$$

Puisque K est un compact, on peut extraire de la suite x'_n une sous-suite convergente $x'_{n_1}, x'_{n_2}, \dots, x'_{n_k}, \dots \rightarrow x_0 \in K$. Comme $\rho(x'_{n_k}, x''_{n_k}) < \frac{1}{n_k}$, on a

$$\rho(x''_{n_k}, x_0) \leq \rho(x'_{n_k}, x''_{n_k}) + \rho(x'_{n_k}, x_0) \rightarrow 0,$$

de sorte que $x'_{n_1}, x'_{n_2}, \dots, x'_{n_k}, \dots \rightarrow x_0$. La fonction $y = f(x)$ est définie et continue pour $x = x_0$; étant donné ε , trouvons $\delta_0 > 0$ tel que

$$\rho[f(x), f(x_0)] < \frac{\varepsilon}{2}$$

dès que $\rho(x, x_0) < \delta_0$. Puisque $x'_{n_k} \rightarrow x_0$, $x''_{n_k} \rightarrow x_0$, la dernière inégalité est satisfaite pour tous les points x'_{n_k}, x''_{n_k} à partir d'un certain numéro. Donc, pour ces points, on a

$$\rho[f(x'_{n_k}), f(x_0)] < \frac{\varepsilon}{2}, \quad \rho[f(x''_{n_k}), f(x_0)] < \frac{\varepsilon}{2},$$

d'où

$$\rho[f(x'_{n_k}), f(x''_{n_k})] < \varepsilon,$$

ce qui contredit (1). Le théorème est démontré.

c. Oscillation d'une fonction. Soit toujours $f(x)$ une fonction définie sur un espace métrique M à valeurs dans un espace métrique P . Construisons la fonction numérique

$$\omega_f(\delta) = \sup_{\rho(x_1, x_2) \leq \delta} \rho(f(x_1), f(x_2)).$$

Elle a pour argument un nombre positif δ . Si la fonction $f(x)$ est uniformément continue sur M , alors $\omega_f(\delta)$ est finie pour tous les δ assez petits et, selon la définition de la continuité uniforme,

$$\lim_{\delta \rightarrow 0} \omega_f(\delta) = 0. \quad (2)$$

En particulier, si M est compact, ce fait a toujours lieu d'après le théorème de Heine.

Réciproquement, si $\omega_f(\delta)$ est finie pour tout $\delta > 0$ assez petit et si l'égalité (2) est satisfaite, alors la fonction $f(x)$ est uniformément continue sur l'espace M .

La fonction $\omega_f(\delta)$ s'appelle *oscillation de la fonction $f(x)$ sur l'espace M* .

d. Pour des fonctions numériques $f(x)$ (qui prennent les valeurs sur la droite numérique R_1 munie de la métrique usuelle $\rho(f, g) = |f - g|$), la fonction $\omega_f(\delta)$ possède les propriétés suivantes :

- 1) $\omega_{f+g}(\delta) \leq \omega_f(\delta) + \omega_g(\delta)$;
- 2) $\omega_{Af}(\delta) = A \omega_f(\delta)$ ($A > 0$ constant) ;
- 3) $\omega_{fg}(\delta) \leq \sup |f(x)| \cdot \omega_g(\delta) + \omega_f(\delta) \cdot \sup |g(x)|$.

Les deux premières propriétés résultent directement de la définition. Pour établir la troisième, on se sert de l'inégalité

$$\begin{aligned} & |f(x_1)g(x_1) - f(x_2)g(x_2)| = \\ & = |f(x_1)g(x_1) - f(x_1)g(x_2) + f(x_1)g(x_2) - \\ & - f(x_2)g(x_2)| \leq |f(x_1)| |g(x_1) - g(x_2)| + \\ & + |f(x_1) - f(x_2)| |g(x_2)| \leq \\ & \leq \sup |f(x)| \cdot |g(x_1) - g(x_2)| + |f(x_1) - f(x_2)| \cdot \sup |g(x)|. \end{aligned}$$

5.18. Fonctions continues de deux variables. En analyse, on a souvent affaire à des fonctions continues de deux variables (ou plus). Soient M_1 et M_2 deux espaces métriques et ρ_1 et ρ_2 les métriques respectives. Une fonction $f(x_1, x_2)$ qui a pour variables indépendantes les points $x_1 \in M_1$ et $x_2 \in M_2$ et pour valeurs les points d'un espace métrique P est dite *continue par rapport à l'ensemble des variables x_1, x_2 pour $x_1 = x_1^0, x_2 = x_2^0$* si, pour tout $\epsilon > 0$, on peut trouver un $\delta > 0$ tel que $\rho[f(x_1, x_2), f(x_1^0, x_2^0)] < \epsilon$ dès que $\rho_1(x_1, x_1^0) < \delta$, $\rho_2(x_2, x_2^0) < \delta$. Il est aisé de voir qu'au fond cette définition n'est pas nouvelle et coïncide avec la définition de la continuité de la fonction $f(x_1, x_2)$ en tant que fonction d'une variable $x = (x_1, x_2)$ qui est un point du produit $M = M_1 \times M_2$ des espaces métriques M_1 et M_2 . Rappelons (3.16) que, dans le produit cartésien M , la métrique est définie suivant la règle : si $x = (x_1, x_2)$, $y = (y_1, y_2)$, on pose $\rho(x, y) = \max[\rho_1(x_1, y_1), \rho_2(x_2, y_2)]$. On peut considérer la fonction $f(x_1, x_2)$ de deux variables comme fonction $f(x)$ du point $x = (x_1, x_2)$; sa continuité en un point $x^0 = (x_1^0, x_2^0)$ signifie que, pour tout $\epsilon > 0$, il existe un $\delta > 0$ tel que $\rho(x, x^0) = \max[\rho_1(x_1, x_1^0), \rho_2(x_2, x_2^0)] < \delta$ implique $|f(x) - f(x^0)| < \epsilon$. L'inégalité $\rho(x, x^0) < \delta$ étant équivalente à deux inégalités $\rho_1(x_1, x_1^0) < \delta$ et $\rho_2(x_2, x_2^0) < \delta$, les deux définitions données de la continuité sont équivalentes.

Nous avons signalé deux autres métriques sur le produit $M = M_1 \times M_2$ (3.16). Vu qu'elles sont homéomorphes à la métrique utilisée ci-dessus (3.34.e), on n'a pas besoin de les expliciter ici ;

il résulte de 5.11.f que l'ensemble des fonctions continues (à valeurs dans un espace fixe P) est le même pour chacune de ces trois métriques.

§ 5.2. Fonctions numériques continues sur la droite numérique

5.21.a. Dans les paragraphes 5.2-5.7, nous considérons des fonctions $f(x)$ d'une variable x parcourant un ensemble E de la droite achevée $\bar{R}$ qui est muni de la métrique $r(x, y)$ de l'espace $\bar{R}$ (3.35.e). Les valeurs de la fonction $f(x)$ sont supposées appartenant elles aussi à l'espace $\bar{R}$ avec la même métrique. D'ailleurs, si x_0 ou bien $f(x_0)$ est fini, alors, la métrique $r(x, y)$ étant homéomorphe à la métrique usuelle $\rho(x, y) = |x - y|$ dans un domaine fini, on peut employer la métrique usuelle (3.13.a) s'il s'agit de la continuité de la fonction $f(x)$ pour $x = x_0$ (respectivement pour les valeurs de x ou de $f(x)$).

b. Considérons de ce point de vue une fonction rationnelle

$$f(x) = \frac{a_0 x^n + a_1 x^{n-1} + \dots + a_n}{b_0 x^m + b_1 x^{m-1} + \dots + b_m}$$

définie pour $x \in R$, sauf pour les points où le dénominateur s'annule. Si l'on complète sa définition par les conditions

$$f(-\infty) = \lim_{x \rightarrow -\infty} f(x), \quad f(\infty) = \lim_{x \rightarrow \infty} f(x)$$

(l'existence de ces limites dans $\bar{R}$ a été démontrée dans 4.39.d), alors on obtient une fonction à valeurs dans $\bar{R}$ qui est aussi bien continue aux points $\pm\infty$.

c. L e m m e. Si une fonction numérique à valeurs dans $\bar{R}$ est continue pour $x = a \in \bar{R}$, alors, quel que soit $\varepsilon > 0$, il existe une boule V_δ centrée en a , i.e. l'ensemble de la forme

$$\begin{aligned} |x - a| &< \delta && \text{si } a \text{ est fini,} \\ x &> \frac{1}{\delta} && \text{si } a = +\infty, \\ x &< -\frac{1}{\delta} && \text{si } a = -\infty, \end{aligned}$$

dont chaque point vérifie l'inégalité:

$$\begin{aligned} |f(x) - f(a)| &< \varepsilon && \text{si } f(a) \text{ est fini,} \\ f(x) &> \frac{1}{\varepsilon} && \text{si } f(a) = +\infty, \\ f(x) &< -\frac{1}{\varepsilon} && \text{si } f(a) = -\infty. \end{aligned}$$

De plus, la boule V_δ peut être choisie de façon qu'en tout son point la valeur de la fonction $f(x)$ soit de même signe que le nombre $f(a)$ si $f(a) \neq 0$.

Ceci résulte de 4.34.a,b, de la définition de la métrique dans $\bar{R}$ (en particulier, 3.35.f) et de la définition de la continuité.

5.22. Théorème (Bolzano). Si une fonction numérique à valeurs dans $\bar{R}$ est continue sur un intervalle $[a, b] \in \bar{R}$ et prend les valeurs de signes contraires à ses extrémités, alors il existe à l'intérieur de l'intervalle un point c pour lequel $f(c) = 0$.

Démonstration. Soit, pour fixer les idées, $f(a) < 0$, $f(b) > 0$. D'après 5.21.c, l'inégalité $f(x) < 0$ reste valable pour tous les x assez proches de a et l'inégalité $f(x) > 0$ l'est pour tous les x assez proches de b . Posons

$$c = \sup \{x \in [a, b] : f(x) < 0\}.$$

Vu ce qui précède, le point c ne peut se confondre avec a , ni avec b . D'après la définition de la borne supérieure, on a $f(x') \geq 0$ pour $x' > c$, et, pour tout $\delta > 0$, il existe un $x'' > c - \delta$ tel que $f(x'') < 0$. Ainsi, dans tout voisinage du point c , il existe des points x' et x'' tels que $f(x') \geq 0$, $f(x'') < 0$. Or, d'après 5.21.c, c'est impossible si $f(c) \neq 0$. Donc $f(c) = 0$, ce qu'il fallait démontrer.

5.23. Conséquence. Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, $A = f(a) \neq f(b) = B$, et si un nombre C est compris entre les nombres A et B , il existe un point $c \in [a, b]$ tel que $f(c) = C$.

En effet, la fonction $f(x)$ étant continue, $f(x) - C$ l'est aussi et vérifie l'hypothèse du théorème de Bolzano. En appliquant le théorème on obtient le résultat voulu.

5.24. Continuité à gauche et à droite. En vue des applications ultérieures, considérons certains types de directions dans le domaine R des nombres réels finis. D'après la définition 4.15.b, la direction $x \rightarrow a$ ($\neq \pm \infty$) est l'ensemble de tous les intervalles $(a - h, a + h)$, $0 < h < h_0$, dépourvus du point a . Définissons la direction $x \nearrow a$ comme ensemble de tous les intervalles $(a - h, a)$ et la direction $x \searrow a$ comme ensemble de tous les intervalles $(a, a + h)$ ($0 < h < h_0$).

Lorsqu'une fonction $f(x)$ est définie dans l'intervalle $(a - h_0, a)$, la question se pose de l'existence de la limite $\lim_{x \nearrow a} f(x)$ que l'on

désigne (si elle existe) par $f(a - 0)$. Lorsqu'une fonction $f(x)$ est définie dans l'intervalle $(a, a + h_0)$, la question se pose de l'existence de la limite $\lim_{x \searrow a} f(x)$ que l'on désigne par $f(a + 0)$.

Si une fonction $f(x)$ est définie dans l'intervalle $(a - h_0, a + h_0)$

(à l'exception éventuelle du point a), la question se pose de l'existence de la limite $\lim_{x \rightarrow a} f(x)$. Etant donné que, pour tout h , $0 < h < h_0$:

$$(a - h, a) \subset (a - h, a + h) - \{a\},$$

$$(a, a + h) \subset (a - h, a + h) - \{a\},$$

$$(a - h, a + h) - \{a\} = (a - h, a) \cup (a, a + h),$$

on peut appliquer les résultats 4.16 en posant

$$E = \{x \in R : x \neq a\}, \quad G = \{x \in R : x < a\},$$

$$H = \{x \in R : x > a\}.$$

Appliqués au cas discuté, ces résultats amènent à la conclusion suivante: l'existence de la limite $\lim_{x \rightarrow a} f(x) = p$ entraîne l'existence des limites $\lim_{x \nearrow a} f(x) = \lim_{x \nearrow a} f(x) = p$, et réciproquement, l'existence et l'égalité des limites $\lim_{x \nearrow a} f(x) = \lim_{x \searrow a} f(x) = p$ entraînent l'existence de la limite $\lim_{x \rightarrow a} f(x) = p$.

Supposons que la fonction $f(x)$ soit également définie pour $x = a$. Nous savons que l'égalité $\lim_{x \rightarrow a} f(x) = f(a)$ définit la classe des fonctions continues pour $x = a$. L'égalité $f(a) = \lim_{x \nearrow a} f(x)$ définit la classe des fonctions continues à gauche pour $x = a$. L'égalité $f(a) = \lim_{x \searrow a} f(x)$ définit la classe des fonctions continues à droite pour $x = a$. Une fonction $f(x)$ continue au point a à gauche et à droite est continue pour $x = a$.

§ 5.3. Fonctions monotones

5.31. Soit $f(x)$ une fonction à valeurs dans $\bar{R}$ définie sur un ensemble $E \subset \bar{R}$. Si $x \in E$, $y \in E$, $x < y$ impliquent l'une des quatre inégalités suivantes:

$$(a) \quad f(x) < f(y),$$

$$(b) \quad f(x) \leq f(y),$$

$$(c) \quad f(x) > f(y),$$

$$(d) \quad f(x) \geq f(y),$$

alors la fonction $f(x)$ est respectivement dite: (a) croissante sur E ; (b) non décroissante sur E ; (c) décroissante sur E ; (d) non croissante sur E . Dans chacun des quatre cas, la fonction $f(x)$ est dite monotone sur E .

5.32. Une fonction croissante (resp. décroissante) ne prend une valeur C que pour une seule valeur de x . Une fonction non croissante (resp. non décroissante) peut prendre une même valeur en plusieurs points.

Les résultats 5.33-5.36 sont formulés et démontrés pour des fonctions non décroissantes; nous laissons au lecteur le soin d'établir les énoncés analogues pour des fonctions non croissantes.

5.33. Supposons que $x_0 \in E$ ne soit pas l'extrémité gauche de l'intervalle $E = [a, b]$ sur lequel une fonction non décroissante $f(x)$ est définie. La fonction $f(x)$ est non décroissante suivant la direction $E(x \nearrow x_0)$ au sens de 4.51; puisqu'elle est majorée sur cette direction (par exemple, par la valeur $f(x_0)$), alors, d'après le théorème 4.53, il existe

$$f(x_0 - 0) = \lim_{E(x \nearrow x_0)} f(x) = \sup \{f(x) : x < x_0, x \in E\}.$$

D'une façon analogue, en considérant la direction $E(x \searrow x_0)$ et en appliquant le théorème 4.54 on aboutit à l'existence de la limite

$$f(x_0 + 0) = \lim_{E(x \searrow x_0)} f(x) = \inf \{f(x) : x > x_0, x \in E\}$$

pour tout point x_0 distinct de l'extrémité droite de l'intervalle $[a, b]$. Pour les points a et b , nous posons par définition :

$$f(a - 0) = f(a), \quad f(b + 0) = f(b).$$

5.34. Puisque $f(x'') \leq f(x_0) \leq f(x')$ quels que soient $x'' \leq x_0$ et $x' \geq x_0$, on a, d'après 1.62, $f(x_0 - 0) \leq f(x_0) \leq f(x_0 + 0)$. Si $f(x_0 - 0) = f(x_0)$, la fonction $f(x)$ est continue à gauche pour $x = x_0$; si $f(x_0) = f(x_0 + 0)$, elle y est continue à droite. L'égalité $f(x_0 - 0) = f(x_0 + 0)$ ainsi que l'égalité $f(x_0 - 0) = f(x_0) = f(x_0 + 0)$ qui en résulte représentent une condition nécessaire et suffisante de continuité d'une fonction non décroissante $f(x)$ pour $x = x_0$.

5.35. Discutons le cas $f(x_0 - 0) < f(x_0 + 0)$. Un tel point x_0 est bien un point de discontinuité de la fonction $f(x)$. Les valeurs de la fonction $f(x)$ à gauche du point x_0 ne dépassent pas $f(x_0 - 0)$, celles à droite du point x_0 sont au moins égales à $f(x_0 + 0)$; la valeur $f(x_0)$ appartient à l'intervalle

$$I = [f(x_0 - 0), f(x_0 + 0)].$$

Or, ceci veut dire qu'aucun nombre de l'intervalle I , à une seule exception éventuelle, n'est une valeur de la fonction $f(x)$. Il en résulte une condition suffisante bien utilisée de continuité d'une fonction non décroissante.

5.36. Si une fonction $f(x)$ non décroissante sur un intervalle $[a, b]$ prend toute valeur de l'intervalle $[f(a), f(b)]$ pour un $x \in [a, b]$, alors $f(x)$ est continue sur $[a, b]$.

En effet, dans l'hypothèse formulée, l'inégalité $f(x_0 - 0) < f(x_0 + 0)$ n'est possible pour aucun point $x_0 \in [a, b]$. Donc, en vertu de 5.34, la fonction $f(x)$ est continue en tout point $x_0 \in [a, b]$.

La condition énoncée est aussi nécessaire (5.23).

5.37. Fonction inverse.

Définition. Soit une fonction $y = f(x)$ définie sur un ensemble E et prenant ses valeurs y dans un ensemble F ; soit une fonction $\varphi(y)$ définie sur l'ensemble F et prenant ses valeurs dans l'ensemble E . La fonction $\varphi(y)$ est dite *inverse* de la fonction $f(x)$ si $\varphi[f(x)] = x$ partout sur E , et $f[\varphi(y)] = y$ partout sur F .

5.38.a. Théorème. Supposons qu'une fonction numérique $y = f(x) \in \bar{R}$ soit définie, continue et croissante sur un intervalle $[a, b] \subset \bar{R}$. Alors, sur l'intervalle $[f(a), f(b)]$, il existe une fonction $\varphi(y)$, inverse de la fonction $f(x)$, qui est aussi continue et croissante.

Démonstration. Pour $y = f(a) = A$, posons $\varphi(A) = a$; pour $y = f(b) = B$, posons $\varphi(B) = b$; pour tout $y = C \in (f(a), f(b))$, posons $\varphi(y)$ égale à la valeur (unique) de x pour laquelle

$$f(x) = C.$$

De cette façon, la fonction $\varphi(y)$ est définie sur $[A, B]$, de sorte que, quels que soient $x \in [a, b]$, $y \in [A, B]$, on a :

$$\varphi[f(x)] = x, \quad f[\varphi(y)] = y.$$

Il est évident que $y_1 < y_2$ implique $x_1 < x_2$, donc $\varphi(y_1) < \varphi(y_2)$, de sorte que la fonction $\varphi(y)$ est croissante. Sa continuité résulte du fait qu'elle prend toutes les valeurs dans l'intervalle $a \leq x \leq b$ (5.36). Le théorème est démontré.

b. Examinons maintenant le cas où la fonction $f(x)$ est donnée non pas sur un intervalle fermé, mais sur un intervalle ouvert, et démontrons qu'il se réduit au précédent. Soit une fonction $y = f(x)$ définie, continue et croissante sur un intervalle $(a, b) \in \bar{R}$. Posons $A = f(a) = \lim_{x \searrow a} f(x)$, $B = f(b) = \lim_{x \nearrow b} f(x)$ (5.33); la fonction

$f(x)$ est alors définie sur l'intervalle $[a, b]$ où elle est toujours continue et croissante. En appliquant le théorème démontré plus haut, nous obtenons : il existe une fonction $\varphi(y)$ définie, continue et croissante sur l'intervalle $[A, B] \subset \bar{R}$ qui est inverse de la fonction $f(x)$. En considérant $f(x)$ sur le seul intervalle (a, b) on aboutit au résultat suivant :

Théorème. Une fonction $y = f(x)$ définie, continue et croissante sur un intervalle $(a, b) \subset \bar{R}$ admet la fonction inverse $\varphi(y)$ définie, continue et croissante sur l'intervalle $(A, B) \in \bar{R}$, où

$$A = \lim_{x \searrow a} f(x), \quad B = \lim_{x \nearrow b} f(x).$$

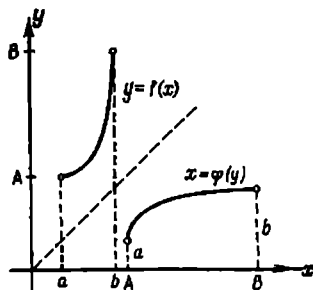


Fig. 5.1.

5.39. Graphique de la fonction inverse. Etant donné le graphique d'une fonction numérique croissante $y = f(x)$ sur un intervalle $[a, b]$, le graphique de la fonction inverse doit, naturellement, être construit sur l'intervalle $[A, B]$, où $A = f(a)$, $B = f(b)$. Si l'on porte ces intervalles sur un même axe, alors le graphique de la fonction inverse

$x = \varphi(y)$ sera symétrique de celui de la fonction $y = f(x)$ par rapport à la bissectrice du premier quadrant: en effet, un point $(x, y = f(x))$ est symétrique du point $(y, x = \varphi(y))$ par rapport à ladite bissectrice (fig. 5.1).

§ 5.4. Logarithme

5.41. **Théorème.** Pour $x > 0$, il existe l'unique fonction $y = f(x)$ satisfaisant aux conditions suivantes:

- 1) $f(xy) = f(x) + f(y)$ quels que soient x, y positifs;
- 2) $f(a) = 1$ pour un $a > 1$ donné;
- 3) $f(x)$ est une fonction croissante: si $x < y$, alors $f(x) < f(y)$.

Démonstration *). Tout d'abord, en supposant que la fonction $f(x)$ existe, trouvons son expression. Il résulte de 1) que $f(1) = 2f(1)$, d'où $f(1) = 0$. Ensuite, il résulte de 1) et 2) que $f(a^n) = nf(a) = n$ pour tout n naturel. D'après 1.72, pour tout $n \geq 1$, on peut trouver un m tel que $a^m \leq x^n < a^{m+1}$. Comme $f(x)$ croît, il résulte de l'inégalité $a^m \leq x^n < a^{m+1}$ que $f(a^m) = m \leq f(x^n) = nf(x) < f(a^{m+1}) = m+1$; ainsi, pour tout $n \geq 1$, on peut trouver un m tel que

$$\frac{m}{n} \leq f(x) < \frac{m+1}{n}.$$

Ceci veut dire que le nombre $f(x)$ est défini d'une façon univoque

*) D'après [17].

comme point commun à tous les intervalles $\left[\frac{m}{n}, \frac{m+1}{n}\right]$ de longueur arbitrairement petite.

A présent, nous utilisons ces considérations pour construire la fonction $f(x)$. Posons $n_1 = 1$, $n_2 = 2$, ..., $n_k = 2^k$ et, pour un k donné, trouvons un m_k de façon à avoir

$$a^{m_k} \leq x^{n_k} < a^{m_k+1}.$$

Ceci détermine le segment $\left[\frac{m_k}{n_k}, \frac{m_k+1}{n_k}\right]$. Montrons que l'intervalle suivant $\left[\frac{m_{k+1}}{n_{k+1}}, \frac{m_{k+1}+1}{n_{k+1}}\right]$ est inclus dans le précédent. En effet, il résulte de l'inégalité

$$a^{m_k} \leq x^{n_k} < a^{m_k+1}$$

que

$$a^{2m_k} \leq x^{2n_k} = x^{n_{k+1}} < a^{2(m_k+1)},$$

d'où $m_{k+1} \geq 2m_k$, $m_{k+1} + 1 \leq 2(m_k + 1)$,

$$\frac{m_k}{n_k} = \frac{2m_k}{2n_k} \leq \frac{m_{k+1}}{n_{k+1}} < \frac{m_{k+1}+1}{n_{k+1}} \leq \frac{2(m_k+1)}{2n_k} = \frac{m_k+1}{n_k}.$$

De la sorte, les intervalles $\left[\frac{m_k}{n_k}, \frac{m_k+1}{n_k}\right]$ forment une suite d'intervalles fermés emboîtés et, d'après le principe de Cantor (1.81), possèdent un point commun; d'après 1.82, il est unique puisque les longueurs des intervalles tendent vers zéro pour $k \rightarrow \infty$. Choisissons pour valeur $f(x)$ le nombre correspondant à ce point commun.

Montrons maintenant que la fonction $f(x)$ ainsi construite satisfait aux conditions 1)-3). Soient x et y ; pour un k donné, trouvons m_k et p_k tels que

$$a^{m_k} \leq x^{n_k} < a^{m_k+1}, \quad a^{p_k} \leq y^{n_k} < a^{p_k+1},$$

de sorte que

$$\frac{m_k}{n_k} \leq f(x) \leq \frac{m_k+1}{n_k}, \quad \frac{p_k}{n_k} \leq f(y) \leq \frac{p_k+1}{n_k}. \quad (1)$$

Il en résulte

$$a^{m_k+p_k} \leq (xy)^{n_k} < a^{m_k+p_k+2}.$$

de sorte que

$$\frac{m_k+p_k}{n_k} \leq f(xy) \leq \frac{m_k+p_k+2}{n_k}.$$

Or, il découle de (1) que

$$\frac{m_k+p_k}{n_k} \leq f(x) + f(y) \leq \frac{m_k+p_k+2}{n_k}.$$

Ainsi

$$|f(x) + f(y) - f(xy)| \leq \frac{2}{n_k};$$

n_k étant arbitraire, nous appliquons 1.57 et aboutissons à l'égalité 1).

Pour $x = a$, nous obtenons l'égalité évidente $f(a) = 1$. Pour tout $x > 1$, il existe un n tel que $a \leq x^n$, donc, selon la définition de la fonction $f(x)$, nous avons $f(x) \geq \frac{1}{n}$. En particulier, pour tout $x > 1$, nous avons $f(x) > 0$. Par conséquent, $f(xy) = f(x) + f(y) > f(x)$ pour tout $y > 1$ et la fonction $f(x)$ est croissante; le théorème est démontré.

5.42. La fonction $f(x)$ qui satisfait aux conditions 1)-3) et dont l'existence et l'unicité viennent d'être démontrées s'appelle *logarithme de x de base a* et se désigne par $\log_a x$. Les formules 1)-3) de 5.41 pour $a > 1$ prennent maintenant la forme :

$$\log_a(xy) = \log_a x + \log_a y; \quad (1)$$

$$\log_a a = 1; \quad (2)$$

$$\text{si } x < y, \text{ alors } \log_a x < \log_a y. \quad (3)$$

Ensuite, (1) implique les formules suivantes :

$$\log_a x^n = n \log_a x \text{ pour tout } n \text{ naturel}; \quad (4)$$

$$\log_a 1 = \log_a 1^2 = 2 \log_a 1, \text{ d'où} \quad (5)$$

$$\log_a 1 = 0;$$

$$\log_a \left(x \cdot \frac{1}{x}\right) = \log_a 1 = \log_a x + \log_a \frac{1}{x} = 0,$$

d'où

$$\log_a \frac{1}{x} = -\log_a x;$$

par conséquent

$$\log_a \frac{x}{y} = \log_a x - \log_a y. \quad (6)$$

Pour cette raison, la formule (4) est valable pour tout n entier.

5.43. Montrons que le *logarithme est une fonction continue de x* . Pour un $\varepsilon > 0$ donné, trouvons $n > \frac{1}{\varepsilon}$ et posons $h = \sqrt[n]{a}$. Comme $h^n = a$, il vient $n \log_a h = \log_a h^n = \log_a a = 1$, de sorte que $\log_a h = \frac{1}{n}$, $\log_a \frac{1}{h} = -\frac{1}{n}$. Donc, dans l'intervalle $\frac{1}{h} \leq x \leq h$, on a l'inégalité

$$-\varepsilon < -\frac{1}{n} \leq \log_a x \leq \frac{1}{n} < \varepsilon,$$

ce qui démontre la continuité de la fonction $\log_a x$ pour $x = 1$.

Pour un y_0 quelconque et pour $y = xy_0$, $\frac{1}{h} < x < h$, nous avons $\log_a y = \log_a x + \log_a y_0$, de sorte que

$$|\log_a y - \log_a y_0| = |\log_a x| \leq \varepsilon.$$

L'intervalle $(\frac{1}{h} y_0, h y_0)$ contenant un intervalle $(y_0 - \delta, y_0 + \delta)$, la fonction $\log_a x$ est continue pour $x = y_0$.

5.44. Établissons la liaison qui existe entre les logarithmes de deux bases différentes a et b . Considérons la fonction

$$f(x) = \frac{\log_a x}{\log_a b}.$$

Il est évident qu'elle satisfait aux conditions 1) et 3) (5.41) qui figurent dans la définition d'un logarithme et, de plus, à l'égalité

$$f(b) = \frac{\log_a b}{\log_a b} = 1.$$

En vertu de l'unicité établie du logarithme, nous avons $f(x) = \log_b x$. Donc, quels que soient $a > 1$, $b > 1$, $x > 0$, on a la formule

$$\log_a x = \log_a b \cdot \log_b x. \quad (1)$$

5.45. Signalons les relations limites

$$\lim_{x \rightarrow \infty} \log_a x = \infty, \quad \lim_{x \searrow 0} \log_a x = -\infty.$$

La première, en vertu du théorème 4.55, résulte du fait que, pour $x \rightarrow \infty$, la fonction $\log_a x$ croît et n'est pas bornée (parce que, par exemple, $\log_a(a^n) = n$ pour tout n). La seconde se ramène à la première d'après l'égalité

$$\log_a \left(x \cdot \frac{1}{x}\right) = \log_a x + \log_a \frac{1}{x} = \log_a 1 = 0.$$

Les relations signalées permettent de compléter la définition d'un logarithme par les conditions $\log_a 0 = -\infty$, $\log_a \infty = \infty$, alors $\log_a x$ s'avère une fonction à valeurs dans $\bar{R}$, continue pour $0 \leq x \leq \infty$.

§ 5.5. Exponentielle

5.51. La fonction continue croissante $y = \log_a x$ admet, d'après le théorème de la fonction inverse (5.38), la fonction inverse qui s'appelle *fonction exponentielle* ou *exponentielle* tout court et qui se désigne par $y = a^x$. Une exponentielle est définie pour tous les x réels, ses valeurs sont des nombres positifs. Ainsi,

$$a^{\log_a x} = x, \quad \log_a(a^x) = x. \quad (1)$$

D'autre part, nous avons pour tout n naturel :

$$\log_a (\underbrace{a \dots a}_{n \text{ fois}}) = \log_a a + \dots + \log_a a = n.$$

La dernière égalité montre que le symbole a^x s'accorde avec le symbole $a^n = a \dots a$ pour un n naturel.

Si $x = 1/n$, nous obtenons $\log_a a^{1/n} = \frac{1}{n}$ d'où

$$n \log_a a^{1/n} = \log_a (a^{1/n})^n = 1, \quad (a^{1/n})^n = a,$$

c'est-à-dire que $a^{1/n}$ est la racine n -ième de a (1.63).

5.52. Tout comme la fonction $y = \log_a x$, la fonction inverse $y = a^x$ est croissante et continue pour tous les x . On a les égalités suivantes :

$$a^{x+y} = a^x \cdot a^y, \quad (1)$$

$$(a^x)^z = a^{xz} \quad (2)$$

quels que soient x, y, z réels. En effet, soient $x = \log_a \xi$, $y = \log_a \eta$, on a

$$a^{x+y} = a^{\log_a \xi + \log_a \eta} = a^{\log_a \xi \eta} = \xi \eta = a^x \cdot a^y;$$

(2) se démontre d'une façon analogue.

En remplaçant x par b^x dans 5.44 (1) on obtient

$$\log_a (b^x) = \log_a b \cdot \log_b b^x = x \log_a b. \quad (3)$$

5.53. Les définitions d'un logarithme et d'une exponentielle peuvent être étendues au cas d'une base $b \in (0, 1)$. Notamment, en posant $b = \frac{1}{a}$, $a > 1$, nous avons par définition (et compte tenu de la formule 5.44 (1)) :

$$\log_b x = \frac{\log_a x}{\log_a b} = -\log_a x = -\log_{1/a} x.$$

Donc, la fonction $\log_b x$ décroît lorsque x croît ; les propriétés analogues à 5.41 1) et 2) restent valables. Ensuite.

$$b^x = 1/a^x = a^{-x};$$

donc, $b^{\log_b x} = a^{-\log_b x} = a^{\log_a x} = x$ et

$$\log_b (b^x) = -\log_a (a^{-x}) = x,$$

de sorte que les formules 5.51 (1), ainsi que les formules 5.52 (1) et (2) qui en découlent, restent valables.

5.54. L'exponentielle est liée d'une façon étroite à la fonction puissance d'un exposant réel quelconque. La fonction puissance est définie pour $x > 0$ selon 5.51, où l'on remplace a par x et x par a :

$$y = x^a.$$

Cette fonction s'exprime par une fonction exponentielle et une fonction logarithmique d'après la formule

$$y = (b^{\log_b x})^a = b^{a \log_b x} \quad (1)$$

pour tout $b < 1$, d'où sa continuité, en vertu du théorème 5.15. La même formule, si l'on y remplace x par pq ($p, q > 0$) et que l'on applique 5.42 (1) et (5.52) (1), amène à l'égalité

$$(pq)^a = p^a \cdot q^a. \quad (2)$$

5.55. Relations limites pour le logarithme et l'exponentielle. Les relations limites suivantes ont lieu :

$$\lim_{x \rightarrow +\infty} b^x = \begin{cases} \infty, & \text{pour } b > 1, \\ 0, & \text{pour } b < 1; \end{cases} \quad (1)$$

$$\lim_{x \rightarrow -\infty} a^x = \begin{cases} 0, & \text{pour } a > 1, \\ \infty, & \text{pour } a < 1; \end{cases} \quad (2)$$

$$\lim_{x \rightarrow \infty} \log_a x = \begin{cases} \infty, & \text{pour } a > 1, \\ -\infty, & \text{pour } a < 1; \end{cases} \quad (3)$$

$$\lim_{x \searrow 0} \log_a x = \begin{cases} -\infty, & \text{pour } a > 1, \\ +\infty, & \text{pour } a < 1; \end{cases} \quad (4)$$

$$\lim_{x \rightarrow \infty} x^b = \begin{cases} \infty, & \text{pour } b > 0, \\ 0, & \text{pour } b < 0; \end{cases} \quad (5)$$

$$\lim_{x \searrow 0} x^b = \begin{cases} 0, & \text{pour } b > 0, \\ \infty, & \text{pour } b < 0. \end{cases} \quad (6)$$

Démonstration. On a la première relation (1) parce que b^x , pour $b > 1$ et $x \rightarrow \infty$, croît et n'est pas bornée (ses valeurs remplissent tout le domaine de définition de la fonction $y = \log_a x$, i.o. toute la droite numérique); la seconde se déduit de la première

en remplaçant b par $\frac{1}{b}$. Les relations (2) s'obtiennent de (1) en remplaçant x par $-x$. La première relation (3) est établie dans 5.45; la seconde s'en obtient en vertu de la définition de $\log_a x$ pour $a < 1$.

En remplaçant dans (3) x par $\frac{1}{x}$ on a (4). Pour déduire (5) et (6), on passe aux logarithmes et on se sert de (1)-(4).

Les relations obtenues permettent de construire les graphiques du logarithme, de l'exponentielle et de la fonction puissance d'exposant quelconque (fig. 5.2-5.4).

La fonction a^x avec $a > 1$ qui est définie pour $x \in \mathbb{R}$ peut l'être aux points $\pm\infty$ en posant :

$$a^{-\infty} = 0, \quad a^{+\infty} = \infty,$$

après quoi elle est continue sur $\bar{\mathbb{R}}$ (à valeurs dans $\bar{\mathbb{R}}$). Si $a < 1$, la dernière affirmation est valable si l'on pose

$$a^{-\infty} = \infty, \quad a^{+\infty} = 0.$$

5.56. Avant de passer à d'autres relations limites, nous démontrons un lemme.

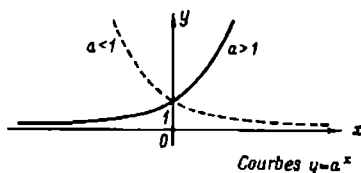


Fig. 5.2.

L e m m e. $\lim_{n \rightarrow \infty} \frac{a^n}{n^r} = \infty$ pour tout $a > 1$ et tout r réel.

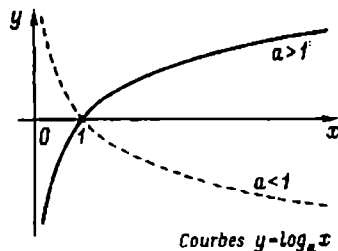


Fig. 5.3.

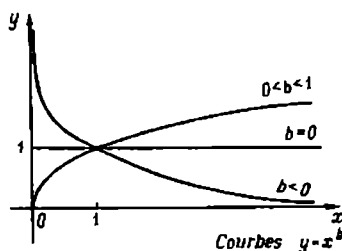


Fig. 5.4.

En effet, soit un nombre b de l'intervalle $(1, a)$. Vu la propriété 5.54 (2), on a l'égalité

$$\frac{a^{n+1}}{(n+1)^r} : \frac{a^n}{n^r} = \frac{a}{\left(1 + \frac{1}{n}\right)^r}. \quad (1)$$

Le second membre de (1) tend vers a pour $n \rightarrow \infty$ en vertu de la continuité de la fonction x^r au point $x = 1$, donc devient, à partir

d'un certain numéro $n = N$, supérieur à b . Alors

$$\begin{aligned}\frac{a^{N+1}}{(N+1)^r} &\geq \frac{a^N}{N^r} \cdot b, \\ \frac{a^{N+2}}{(N+2)^r} &\geq \frac{a^{N+1}}{(N+1)^r} \cdot b \geq \frac{a^N}{N^r} \cdot b^2, \\ &\dots \dots \dots \\ \frac{a^{N+k}}{(N+k)^r} &\geq \frac{a^N}{N^r} \cdot b^k,\end{aligned}$$

et l'égalité

$$\lim_{k \rightarrow \infty} b^k = \infty$$

(qui résulte de 5.55 (1)) démontre le lemme.

5.57. Le lemme 5.56 étant démontré, signalons encore quelques relations limites pour le logarithme et l'exponentielle.

T h é o r è m e. *Les relations limites suivantes sont valables :*

$$\lim_{x \rightarrow \infty} \frac{a^x}{x^r} = \infty ; \quad (1)$$

$$\lim_{x \rightarrow \infty} \frac{\log_a x}{x^p} = 0 ; \quad (2)$$

$$\lim_{x \searrow 0} x^p \log_a x = 0 ; \quad (3)$$

$$\lim_{x \searrow 0} x^{(x^p)} = 1 ; \quad (4)$$

$$\lim_{x \rightarrow \infty} x^{1/x^p} = 1. \quad (5)$$

Ici $a > 1$, r est un nombre réel quelconque et p un nombre positif quelconque.

D é m o n s t r a t i o n. Si l'on trouve, pour un $x > 0$ donné, un n tel que $n \leq x < n + 1$, alors

$$\frac{a^x}{x^r} \geq \frac{a^n}{(n+1)^r} = \frac{1}{a} \frac{a^{n+1}}{(n+1)^r},$$

d'où (1) on vertu du lemme 5.56. La relation (2) s'obtient de (1) en remplaçant x par $\log_a x$ et en élevant à la puissance $p = \frac{1}{r}$.

En remplaçant x par $\frac{1}{x}$ dans (2), on trouve (3). Enfin, (4) résulte de (3) en se débarrassant des logarithmes, et (5) de (4) en remplaçant x par $\frac{1}{x}$. Ainsi, l'exponentielle $y = a^x$, $a > 1$, croît, pour $x \rightarrow \infty$,

plus vite que toute puissance x^r . La fonction logarithmique

$$y = \log_b x \quad (b > 1),$$

au contraire, croît, pour $x \rightarrow \infty$, moins vite que toute puissance x^β ($\beta > 0$).

5.58. Compte tenu du lien qui existe entre la limite d'une fonction $f(x)$ pour $x \rightarrow \infty$ et la limite d'une suite $f(n)$ pour $n \rightarrow \infty$ (4.65), nous pouvons trouver les limites suivantes :

$$\lim_{n \rightarrow \infty} b^n = \begin{cases} \infty, & \text{pour } b > 1 \\ 0, & \text{pour } b < 1, \end{cases} \quad (1)$$

(ceci résulte de 5.55 (1)) ;

$$\lim_{n \rightarrow \infty} \log_a n = \begin{cases} \infty, & \text{pour } a > 1, \\ -\infty, & \text{pour } a < 1 \end{cases} \quad (2)$$

(ceci résulte de 5.55 (3)) ;

$$\lim_{n \rightarrow \infty} n^b = \begin{cases} \infty, & \text{pour } b > 0, \\ 0, & \text{pour } b < 0 \end{cases} \quad (3)$$

(ceci résulte de 5.55 (5)) ;

$$\lim_{n \rightarrow \infty} \frac{a^n}{n^r} = \infty \quad (a > 1, -\infty < r < \infty) \quad (4)$$

(ceci résulte de 5.57 (1)) ;

$$\lim_{n \rightarrow \infty} \frac{\log_a n}{n^p} = 0 \quad (a > 1, p > 0) \quad (5)$$

(ceci résulte de 5.57 (2)) ;

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1 \quad (6)$$

(ceci résulte de 5.57 (5) pour $p = 1$).

5.59. La fonction exponentielle et le nombre e . D'après 4.63.d,

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e = 2,71828 \dots \quad (1)$$

a. Considérons la fonction $f(x) = \left(1 + \frac{1}{x}\right)^x$ définie pour $x \neq 0$. Il se trouve aussi que

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e. \quad (2)$$

Soulignons que ceci ne découle pas directement (4.65) de la relation (1), il faut utiliser les particularités de la fonction $f(x)$. Pour un $x > 1$ donné, trouvons un $n = n(x)$ naturel tel que $n \leq x < n + 1$; alors $n \rightarrow \infty$ lorsque $x \rightarrow \infty$. Nous avons

$$\left(1 + \frac{1}{x}\right)^x \leq \left(1 + \frac{1}{n}\right)^{n+1} = \left(1 + \frac{1}{n}\right)^n \left(1 + \frac{1}{n}\right),$$

$$\left(1 + \frac{1}{x}\right)^x \geq \left(1 + \frac{1}{n+1}\right)^n = \left(1 + \frac{1}{n+1}\right)^{n+1} \cdot \frac{1}{1 + \frac{1}{n+1}}.$$

Les seconds membres de ces deux inégalités ont la même limite e pour $n \rightarrow \infty$. Pour un $\varepsilon > 0$ donné, trouvons un N tel que, pour tout $n > N$, la différence entre chacun de ces membres et le nombre e soit inférieure à ε . Alors, pour $x > N + 1$, la quantité intermédiaire $\left(1 + \frac{1}{x}\right)^x$ différera de e d'un nombre inférieur à ε . Ceci veut dire que (2) est établi.

b. Montrons encore que

$$\lim_{x \rightarrow -\infty} \left(1 + \frac{1}{x}\right)^x = e. \quad (3)$$

En effet, nous avons pour $y \rightarrow \infty$:

$$\begin{aligned} \left(1 - \frac{1}{y}\right)^{-y} &= \left(\frac{y-1}{y}\right)^{-y} = \frac{1}{\left(\frac{y-1}{y}\right)^y} = \left(\frac{y}{y-1}\right)^y = \\ &= \left(1 + \frac{1}{y-1}\right)^{y-1} \left(1 + \frac{1}{y-1}\right) \rightarrow e, \end{aligned}$$

d'où, en remplaçant y par $-x$ et appliquant 5.59.a et 4.17.c, nous obtenons le résultat cherché.

c. En particulier, il découle de b que

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}.$$

C'est un résultat que l'on aurait pu obtenir directement de la définition de e (1).

d. T h é o r è m e.

$$\lim_{z \rightarrow \infty} \frac{\log_a(1+z)}{z} = \log_a e. \quad (4)$$

D é m o n s t r a t i o n. Nous avons $\frac{\log_a(1+z)}{z} = \log_a(1+z)^{\frac{1}{z}}$; posons $z = \frac{1}{x}$ et, en nous basant sur a, b et sur la continuité du logarithme, nous arrivons au résultat cherché.

La formule (4) prend une forme particulièrement simple si l'on choisit pour base a le nombre e . Les logarithmes de base e sont appelés *logarithmes naturels* (ou *népériens*) et désignés comme suit:

$$\log_e x = \log \text{ nat } x = \ln x.$$

En cas de logarithmes naturels, la formule (4) se met sous la forme

$$\lim_{z \rightarrow 0} \frac{\ln(1+z)}{z} = 1. \quad (5)$$

En employant l'unité asymptotique E (4.39), ce résultat prend la forme suivante:

$$\ln(1+z) = zE \quad (E \rightarrow 1 \text{ pour } z \rightarrow 0). \quad (6)$$

e. T h é o r è m e.

$$\lim_{h \rightarrow 0} \frac{a^h - 1}{h} = \frac{1}{\log_a a}. \quad (7)$$

Démonstration. Posons $a^h - 1 = \frac{1}{x}$; alors $a^h = 1 + \frac{1}{x}$, $h = \log_a \left(1 + \frac{1}{x}\right)$ et

$$\frac{a^h - 1}{h} = \frac{\frac{1}{x}}{\log_a \left(1 + \frac{1}{x}\right)} = \frac{1}{\log_a \left(1 + \frac{1}{x}\right)^x},$$

d'où, en vertu de a et b, de la continuité du logarithme et de la correspondance des directions $h \rightarrow 0$ et $x \rightarrow \pm\infty$, il découle le résultat cherché (4.17.a). Avec des logarithmes naturels, on peut le mettre sous une forme plus simple:

$$\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1. \quad (8)$$

La même formule avec l'unité asymptotique est:

$$e^h = 1 + Eh, \quad E \rightarrow 1 \text{ pour } h \rightarrow 0. \quad (9)$$

f. Notons, pour le théorème suivant, que bien que le nombre e ne figure pas dans son énoncé, il serait difficile de le démontrer sans employer ce nombre aussi que des limites qui y sont liées.

T h é o r è m e. Pour tout p réel,

$$\lim_{z \rightarrow 0} \frac{(1+z)^p - 1}{z} = p. \quad (10)$$

Démonstration. En appliquant (6) et (9) on trouve:

$$\begin{aligned} \ln(1+z) &= zE_1, \quad 1+z = e^{zE_1}; \\ (1+z)^p &= e^{pzE_1} = 1 + pzE_1E_2 = 1 + pzE, \end{aligned} \quad (11)$$

ce qu'il fallait démontrer.

La relation (11) montre, entre autres, que le résultat 4.39.b démontré pour un p naturel est valable pour tout p réel.

g. Insistons sur la relation (11) pour $p = -1$:

$$\frac{1}{1+z} = 1 - zE. \quad (12)$$

Voici une relation plus générale et assez utilisée: pour $\alpha \neq \beta$

$$\frac{1+\alpha z E_1}{1+\beta z E_2} = 1 + (\alpha - \beta) z E \quad (E_1, E_2, E \rightarrow 1 \text{ pour } z \rightarrow 0). \quad (13)$$

Elle se déduit de (12) à l'aide de simples transformations suivantes:

$$\begin{aligned} \frac{1+\alpha z E_1}{1+\beta z E_2} &= (1 + \alpha z E_1) (1 - \beta z E_2 E_3) = \\ &= 1 + \alpha z E_1 - \beta z E_2 E_3 - \alpha \beta z^2 E_1 E_2 E_3 = \\ &= 1 + z (\alpha E_1 - \beta E_2 E_3 - \alpha \beta z E_1 E_2 E_3); \end{aligned}$$

la dernière expression entre parenthèses ayant, évidemment, $\alpha - \beta$ pour limite, on obtient définitivement (13).

§ 5.6. Fonctions trigonométriques

5.61. Ce n'est pas sur les axiomes des nombres réels que sont basées les définitions des fonctions trigonométriques employées en géométrie, et nous n'insisterons pas sur ces définitions. En trigonométrie, on déduit les relations:

$$\sin^2 x + \cos^2 x = 1; \quad (1)$$

$$\sin(x + y) = \sin x \cos y + \sin y \cos x; \quad (2)$$

$$\cos(x + y) = \cos x \cos y - \sin x \sin y; \quad (3)$$

$$0 < \sin x < x < \frac{\sin x}{\cos x} \quad (4)$$

pour tout $x > 0$ assez petit, par exemple pour $0 < x < \varepsilon_0$.

Nous *définissons* $\sin x$ et $\cos x$ comme fonctions données pour tous les x réels et vérifiant les relations (1)-(4). Plus loin, dans le chapitre 8, nous donnerons une démonstration de l'existence et de l'unicité de ces fonctions; les raisonnements qui suivent sont, à strictement parler, formelles (s'il existe des fonctions $\sin x$ et $\cos x$ vérifiant (1)-(4), alors . . .).

5.62. En posant dans 5.61 (1)-(3) $x = y = 0$, nous avons les égalités suivantes:

$$\sin^2 0 + \cos^2 0 = 1; \quad (1)$$

$$\sin 0 = 2 \sin 0 \cdot \cos 0; \quad (2)$$

$$\cos 0 = \cos^2 0 - \sin^2 0. \quad (3)$$

Il résulte de (2): ou bien $\sin 0 = 0$, ou bien $\cos 0 = \frac{1}{2}$. Or, (1) et (3) impliquent

$$1 + \cos 0 = 2 \cos^2 0, \quad (4)$$

d'où l'on voit que l'égalité $\cos 0 = \frac{1}{2}$ est impossible. Donc $\sin 0 = 0$, et il découle de (1) que $\cos 0 = \pm 1$. L'égalité (4) permet de choisir de ces deux valeurs de $\cos 0$ la seule possible $\cos 0 = 1$. Donc

$$\sin 0 = 0, \quad \cos 0 = 1. \quad (5)$$

5.63. En remplaçant dans 5.61 (2), (3) y par $-x$ et en appliquant 5.62 (5), nous avons

$$0 = \sin x \cos (-x) + \sin (-x) \cos x,$$

$$1 = \cos x \cos (-x) - \sin x \sin (-x).$$

La résolution de ce système linéaire par rapport à $\sin (-x)$ et $\cos (-x)$ donne

$$\sin (-x) = -\sin x, \quad \cos (-x) = \cos x. \quad (1)$$

En soustrayant de l'égalité 5.61 (2) la même égalité où y est remplacé par $-y$ et en appliquant (1) on trouve

$$\sin (x + y) - \sin (x - y) = 2 \sin y \cos x.$$

Le même procédé appliqué à l'égalité 5.61 (3) aboutit à

$$\cos (x + y) - \cos (x - y) = -2 \sin x \sin y.$$

En remplaçant dans les formules obtenues $x + y$ par α et $x - y$ par β , de sorte que $x = \frac{\alpha + \beta}{2}$, $y = \frac{\alpha - \beta}{2}$, mettons ces formules sous la forme

$$\sin \alpha - \sin \beta = 2 \sin \frac{\alpha - \beta}{2} \cos \frac{\alpha + \beta}{2}; \quad (2)$$

$$\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}. \quad (3)$$

5.64.a. Il découle de 5.61 (1) que, pour tout x ,

$$|\sin x| \leq 1, \quad |\cos x| \leq 1.$$

b. En appliquant a et 5.63 (2), (3) nous démontrerons la continuité des fonctions $\sin x$ et $\cos x$. Supposons que $x > y$ et que la différence $x - y$ soit assez petite. Alors

$$\begin{aligned} |\sin x - \sin y| &= 2 \sin \frac{x - y}{2} \left| \cos \frac{x + y}{2} \right| \leq \\ &\leq 2 \sin \frac{x - y}{2} < 2 \cdot \frac{x - y}{2} = x - y, \end{aligned}$$

d'où la continuité de la fonction $\sin x$. D'une façon analogue,

$$|\cos x - \cos y| = 2 \left| \sin \frac{x + y}{2} \right| \sin \frac{x - y}{2} \leq 2 \sin \frac{x - y}{2} \leq x - y,$$

d'où la continuité de la fonction $\cos x$.

c. D'après la relation 5.61 (4) et compte tenu du théorème 4.34.h, on a

$$\lim_{x \searrow 0} \frac{\sin x}{x} = 1, \quad (1)$$

parce que 5.61 (4) implique, pour $0 < x < \varepsilon_0$,

$$\frac{1}{\cos x} < \frac{\sin x}{x} < 1,$$

et les membres extrêmes de cette inégalité tendent vers 1 pour $x \searrow 0$. De plus, en tenant compte de 5.63 (1) on a

$$\lim_{x \nearrow 0} \frac{\sin x}{x} = \lim_{(-x) \searrow 0} \frac{\sin(-x)}{-x} = 1. \quad (2)$$

En réunissant les formules (1) et (2) on obtient

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \quad (3)$$

5.65.a. Montrons maintenant que la fonction $y = \cos x$ s'annule pour certains $x > 0$. Posons $\beta = \inf_{x > 0} \{\cos x\}$. Si $\beta > 0$, alors il résulte de 5.63 (2) que, pour un x assez petit et pour tout $m = 1, 2, \dots$,

$$\sin(m+1)x - \sin mx > 2\beta \sin \frac{x}{2} > 0;$$

donc les valeurs $\sin mx$ croissent indéfiniment avec m , ce qui contredit 5.64.a. Ainsi $\beta = 0$. Montrons qu'il existe un point x_0 tel que $\cos x_0 = 0$. Supposons que $\cos x > 0$ pour tous les $x \geq 0$. Il résulte alors de 5.63 (2) que la fonction $\sin x$ croît, donc (d'après 5.61 (1)) $\cos x$ décroît. Etant donné $\inf_{x > 0} \{\cos x\} = 0$, il existe un

point x_0 tel que $\cos x < \frac{5}{13}$ pour $x > x_0$ et, par conséquent, $\sin x = \sqrt{1 - \cos^2 x} > \frac{12}{13}$. Pour $y = x > x_0$, la formule 5.61 (2) fournit

$$\frac{12}{13} < \sin 2x = 2 \sin x \cdot \cos x < 2 \cdot 1 \cdot \frac{5}{13} = \frac{10}{13},$$

ce qui est impossible. Donc, il existe des points où $\cos x$ s'annule.

b. Posons

$$\frac{\pi}{2} = \inf \{x > 0 : \cos x = 0\}. \quad (1)$$

Le nombre $\pi/180$ s'appelle *degré*; ainsi $\pi/2$ vaut 90 degrés (90°). La fonction $\cos x$ étant continue, $\cos \frac{\pi}{2} = 0$, $\sin \frac{\pi}{2} = 1$. Sur l'intervalle $[0, \frac{\pi}{2}]$ la fonction $\sin x$ croît et $\cos x$ décroît. Ensuite, les

formules 5.61 (2), (3) donnent

$$\sin\left(\frac{\pi}{2} + x\right) = \sin\frac{\pi}{2} \cos x + \cos\frac{\pi}{2} \sin x = \cos x, \quad (2)$$

$$\cos\left(\frac{\pi}{2} + x\right) = \cos\frac{\pi}{2} \cos x - \sin\frac{\pi}{2} \sin x = -\sin x, \quad (3)$$

d'où

$$\sin(x + \pi) = \cos\left(x + \frac{\pi}{2}\right) = -\sin x, \quad (4)$$

$$\cos(x + \pi) = -\sin\left(x + \frac{\pi}{2}\right) = -\cos x. \quad (5)$$

c. Ensuite,

$$\sin(x + 2\pi) = -\sin(x + \pi) = \sin x, \quad (6)$$

$$\cos(x + 2\pi) = -\cos(x + \pi) = \cos x. \quad (7)$$

Une fonction $y = f(x)$ qui satisfait, pour un T et pour tous les x , à l'équation $f(x + T) = f(x)$ est dite *périodique* et le nombre T

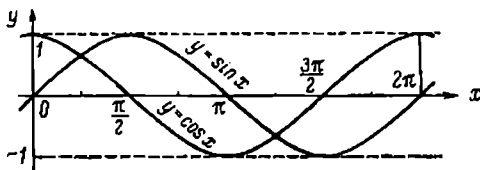


Fig. 5.5.

s'appelle sa *période*. Nous voyons que les fonctions $\sin x$ et $\cos x$ sont des fonctions périodiques de période 2π . Pour une période 2π , les signes de ces fonctions déterminés d'après les formules (2)-(5) varient de la façon suivante

$$\left. \begin{array}{c|c|c|c} 0 < x < \frac{\pi}{2} & \frac{\pi}{2} < x < \pi & \pi < x < \frac{3\pi}{2} & \frac{3\pi}{2} < x < 2\pi \\ \sin x & + & + & - & - \\ \cos x & + & - & - & + \end{array} \right\} \quad (8)$$

Nous voyons que si $\sin x \neq 0$ et $\cos x \neq 0$ pour un x donné, alors les signes de $\sin x$ et $\cos x$ déterminent d'une manière univoque le quart de l'intervalle $0 < x < 2\pi$ qui contient x .

Les graphiques des fonctions $\sin x$ et $\cos x$ sont représentés sur la fig. 5.5. (Pour le moment, nous ne sommes pas en mesure de prouver que le sens de la convexité des courbes est tel que l'on voit sur les graphiques; cela sera fait en 7.55.) Le nombre π vaut 3,14159... (exercice 15 pour le chapitre 9).

5.66. L e m m e. Si, pour certains x et u , les égalités

$$\left. \begin{aligned} \sin(x+u) &= \sin u, \\ \cos(x+u) &= \cos u \end{aligned} \right\} \quad (1)$$

ont lieu, alors $x = 2\pi k$, où k est un nombre entier.

D é m o n s t r a t i o n. En résolvant le système d'équations

$$\sin(x+u) = \sin x \cos u + \cos x \sin u = \sin u,$$

$$\cos(x+u) = \cos x \cos u - \sin x \sin u = \cos u$$

par rapport à $\cos x$ et $\sin x$ on trouve

$$\cos x = 1, \quad \sin x = 0. \quad (2)$$

Dans l'intervalle $0 \leq x < 2\pi$, d'après 5.65, les équations (2) ne sont satisfaites que par $x = 0$. Comme les fonctions $\cos x$ et $\sin x$ ont 2π pour période, les équations (2) sont satisfaites par $x = 2k\pi$ ($k = 0, \pm 1, \pm 2, \dots$) et ne le sont par aucune autre valeur de x . Le lemme est démontré.

5.67. Nous avons vu que la fonction $\sin x$ croît sur l'intervalle $\left[0, \frac{\pi}{2}\right]$; en vertu de l'égalité $\sin(-x) = -\sin x$, elle croît même sur l'intervalle $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ en variant de -1 à $+1$. D'après le théorème de la fonction inverse (5.38), sa fonction inverse est définie sur l'intervalle $[-1, 1]$ et se note $\arcsin x$; cette fonction est croissante, continue et prend ses valeurs dans l'intervalle $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

La fonction $\cos x = \sin\left(x + \frac{\pi}{2}\right)$ croît sur l'intervalle $[-\pi, 0]$ en variant de -1 à $+1$; donc sa fonction inverse que l'on désigne par $\arccos x$ est définie sur l'intervalle $[-1, 1]$. Cette fonction est croissante, continue et prend ses valeurs dans l'intervalle $[-\pi, 0]$.

Etant donné que, d'après 5.65 (2),

$$\sin\left(x + \frac{\pi}{2}\right) = \cos x,$$

nous désignons par u la valeur commune de ces deux quantités et notons

$$x = \arccos u, \quad x + \frac{\pi}{2} = \arcsin u,$$

par conséquent, pour tout $u \in [-1, 1]$

$$\arccos u + \frac{\pi}{2} = \arcsin u. \quad (1)$$

On peut définir une autre fonction $\arccos u$ en partant de la branche décroissante de la fonction $\cos x$ sur l'intervalle $[0, \pi]$.

La fonction inverse est décroissante elle aussi sur l'intervalle $[-1, 1]$ et varie de π à 0. Au lieu de la formule $\sin\left(x + \frac{\pi}{2}\right) = \cos x$ liant les fonctions $\sin x$ et $\cos x$ sur les intervalles où l'on construisait la fonction inverse, il faut cette fois utiliser la formule $\sin\left(\frac{\pi}{2} - x\right) = \cos x$, d'où, à la place de la formule (1), on obtient

$$\arccos u = \frac{\pi}{2} - \arcsin u. \quad (2)$$

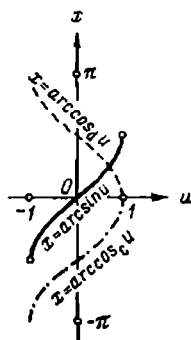


Fig. 5.8.

Pour éviter toute confusion on peut désigner la fonction inverse croissante par $\arccos_c u$ et la fonction inverse décroissante par $\arccos_d u$. Les graphiques des fonctions $\arcsin u$, $\arccos_c u$, $\arccos_d u$ sont représentés fig. 5.6.

Il existe d'autres fonctions sur $[-1, 1]$ qu'on pourrait appeler fonctions inverses de $\sin x$ ou $\cos x$, mais que nous ne considérons pas.

5.68. Posons maintenant

$$\operatorname{tg} x = \frac{\sin x}{\cos x}.$$

Cette fonction est définie partout, sauf les points où $\cos x = 0$,

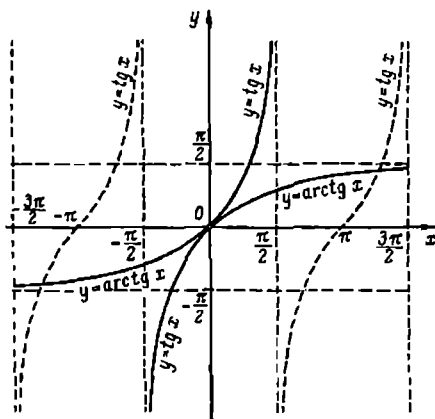


Fig. 5.7.

i.e. les points $x = \frac{\pi}{2} + k\pi$, $k = 0, \pm 1, \pm 2, \dots$. De plus,

$$\lim_{x \nearrow \frac{\pi}{2}} \operatorname{tg} x = +\infty, \quad \lim_{x \searrow -\frac{\pi}{2}} \operatorname{tg} x = -\infty.$$

Il résulte des égalités 5.65 (4), (5) que la fonction $\operatorname{tg} x$ a π pour période. Le comportement des fonctions $\sin x$ et $\cos x$ sur l'intervalle $(-\frac{\pi}{2}, \frac{\pi}{2})$ montre que $\operatorname{tg} x$ croît dans cet intervalle en remplissant par ses valeurs toute la droite numérique. D'après le théorème de la fonction inverse, il existe la fonction inverse $\operatorname{arctg} x$ définie sur toute la droite numérique, croissante, continue et variant dans l'intervalle $(-\frac{\pi}{2}, \frac{\pi}{2})$. Son graphique est représenté fig. 5.7.

Nous laissons au lecteur le soin de considérer les autres fonctions trigonométriques $\operatorname{cotg} x = \frac{1}{\operatorname{tg} x}$, $\sec x = \frac{1}{\cos x}$, $\operatorname{cosec} x = \frac{1}{\sin x}$ ainsi que leurs inverses.

§ 5.7. Applications des fonctions trigonométriques

5.71. Coordonnées polaires dans le plan. Pour tout point (x, y) dans le plan réel R_2 avec $x^2 + y^2 \neq 0$, on a l'égalité

$$\frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2} = 1.$$

Définissons le nombre θ de l'intervalle $0 \leq \theta < 2\pi$ avec les conditions

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}, \quad \sin \theta = \frac{y}{\sqrt{x^2 + y^2}}. \quad (1)$$

Le θ cherché existe et est unique. Pour $x > 0$, $y > 0$, cela résulte de la formule 5.61 (1), de la continuité et de la monotonie des fonctions $\cos \theta$ et $\sin \theta$ pour $0 < \theta < \frac{\pi}{2}$ et de la conséquence du théorème de Bolzano 5.23, dans les autres cas, des formules 5.65 (2)-(5) et de la règle des signes 5.65 (8). Le nombre θ est appelé *angle polaire* du point (x, y) . Il vaut 0 pour les points avec $x > 0$, $y = 0$, $\frac{\pi}{2} = 90^\circ$ pour les points avec $x = 0$, $y > 0$, $\pi = 180^\circ$ pour les points avec $x < 0$, $y = 0$ et $\frac{3\pi}{2} = 270^\circ$ pour les points avec $x = 0$, $y < 0$. D'ailleurs, il est parfois plus commode, en se servant de la périodicité des fonctions $\sin \theta$ et $\cos \theta$, de sous-entendre par θ non pas la seule solution du système (1) dans $[0, 2\pi]$, mais aussi tout nombre réel distinct de cette solution d'un multiple de 2π . Le

nombre

$$r = \sqrt{x^2 + y^2} \quad (2)$$

s'appelle *rayon polaire* du point (x, y) . Les formules (1) peuvent être mises sous la forme

$$x = r \cos \theta, \quad y = r \sin \theta. \quad (3)$$

Les nombres r et θ sont appelés *coordonnées polaires* du point (x, y) ; ils sont définis partout, sauf le point $x = 0, y = 0$ où $r = 0$ et la grandeur θ perd le sens. La ligne $r = \text{const}$ est une circonférence

$$x^2 + y^2 = r^2;$$

la ligne $\theta = \text{const}$ est une demi-droite

$$\frac{y}{x} = \operatorname{tg} \theta.$$

5.72. Représentation trigonométrique des nombres complexes. Soit un nombre complexe $z = x + iy \neq 0$ (§ 2.7). Avec les égalités 5.71 (3), on peut mettre ce nombre sous la forme

$$z = r (\cos \theta + i \sin \theta), \quad (1)$$

θ et r étant définis d'après les formules 5.71 (1), (2). Dans cette notation, r est appelé *module* et θ *argument* du nombre complexe z *). On note $r = |z|$, $\theta = \arg z$.

Si

$$z_1 = r_1 (\cos \theta_1 + i \sin \theta_1),$$

$$z_2 = r_2 (\cos \theta_2 + i \sin \theta_2),$$

alors, en multipliant ces égalités et en appliquant 5.61 (2), (3) on obtient

$$\begin{aligned} z_1 z_2 &= r_1 r_2 [(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) + \\ &\quad + i (\cos \theta_1 \sin \theta_2 + \sin \theta_1 \cos \theta_2)] = \\ &= r_1 r_2 [\cos (\theta_1 + \theta_2) + i \sin (\theta_1 + \theta_2)]. \end{aligned} \quad (2)$$

Donc, en multipliant deux nombres complexes, leurs modules se multiplient et leurs arguments s'additionnent.

5.73. En particulier, pour tout $z \neq 0$, il découle de 5.72 (2) que, quel que soit $m = 1, 2, \dots$ naturel,

$$z^m = r^m (\cos m\theta + i \sin m\theta). \quad (1)$$

En nous servant de la formule (1), nous résolvons l'équation

$$z^m = w, \quad (2)$$

*) Ne pas confondre avec l'argument d'une fonction (2.81).

où w est un nombre complexe donné. Pour $w = 0$, elle n'a, évidemment, que la solution unique $z = 0$. Supposons donc que $w \neq 0$,

$$w = R (\cos \omega + i \sin \omega).$$

Cherchons z sous la forme 5.72 (1), où r et θ sont à déterminer. L'égalité (1) amène à l'équation

$$r^m (\cos m\theta + i \sin m\theta) = R (\cos \omega + i \sin \omega),$$

d'où

$$r^m \cos m\theta = R \cos \omega, \quad r^m \sin m\theta = R \sin \omega. \quad (3)$$

En élevant ces égalités au carré et en additionnant, trouvons

$$r^{2m} = R^2;$$

cette équation possède, dans le domaine des nombres positifs, l'unique solution

$$r = \sqrt[m]{R}.$$

En divisant maintenant les égalités (3) par $r^m = R$, nous avons

$$\cos m\theta = \cos \omega, \quad \sin m\theta = \sin \omega.$$

En vertu de 5.66, où l'on pose $u = \omega$, $x + u = m\theta$, de sorte que $x = m\theta - \omega$, nous avons $m\theta - \omega = 2k\pi$, c'est-à-dire que

$$0 = \frac{\omega}{m} + \frac{2k\pi}{m}$$

$$(k = 0, \pm 1, \pm 2, \dots).$$

Pour $k = 0, 1, \dots, m-1$, on en tire m solutions différentes de l'équation (2):

$$z_k = \sqrt[m]{R} \left[\cos \left(\frac{\omega}{m} + \frac{2k\pi}{m} \right) + \right.$$

$$\left. + i \sin \left(\frac{\omega}{m} + \frac{2k\pi}{m} \right) \right]. \quad (4) \quad \text{Fig. 5.8. } m = 6, \quad R = 1, \quad \omega = 192^\circ$$

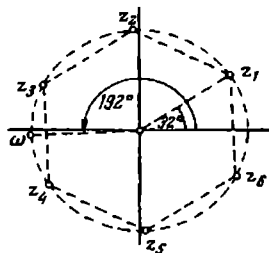
Les autres valeurs de k donnent les valeurs de θ distinctes de celles déjà trouvées d'un multiple de 2π . Ceci ne donne aucune nouvelle racine de l'équation (2).

Les racines trouvées (4) sont disposées aux sommets d'un m -gone régulier centré à l'origine des coordonnées (fig. 5.8.).

5.74. Angle de deux vecteurs dans un espace n -dimensionnel. Etant donnés deux vecteurs

$$x = (x_1, \dots, x_n) \in R_n,$$

$$y = (y_1, \dots, y_n) \in R_n,$$



nous formons leur produit scalaire (2.68 (1))

$$(x, y) = \sum_{k=1}^n x_k y_k.$$

Compte tenu de l'inégalité de Cauchy 2.68 (3):

$$|(x, y)| \leq |x| \cdot |y|.$$

nous pouvons trouver, pour $x \neq 0$, $y \neq 0$, un nombre ω , $0 \leq \omega \leq \pi$, à partir de la condition

$$\cos \omega = \frac{(x, y)}{|x| \cdot |y|}.$$

Le nombre ω s'appelle *angle des vecteurs* x et y et se note $\widehat{(x, y)}$. Si $\omega = 0$, c'est-à-dire que

$$(x, y) = |x| \cdot |y|,$$

alors les vecteurs x et y sont *linéairement dépendants*. En effet, l'équation

$$\begin{aligned} 0 &= (x - \lambda y, x - \lambda y) = (x, x) - 2\lambda (x, y) + \lambda^2 (y, y) = \\ &= |x|^2 - 2\lambda |x| \cdot |y| + \lambda^2 |y|^2 = (|x| - \lambda |y|)^2 \end{aligned}$$

admet, dans ce cas, la solution $\lambda = \frac{|x|}{|y|}$; donc, avec cette valeur de λ , nous avons

$$x = \lambda y.$$

D'une façon analogue, si $\omega = \pi$, c'est-à-dire que $(x, y) = -|x| \times |y|$, les vecteurs x et y sont linéairement dépendants; le facteur de proportionnalité est, cette fois, négatif.

Si $\omega = \frac{\pi}{2}$, donc $(x, y) = 0$, les vecteurs x et y sont dits *orthogonaux*. Cette dernière définition s'étend, naturellement, au cas où l'un des vecteurs x , y (ou les deux) est nul; le vecteur nul est orthogonal à n'importe quel vecteur x .

Appliquons la définition formulée à deux vecteurs du plan. Soient $x \in R_2$ un vecteur de coordonnées polaires ρ , φ et y un vecteur de coordonnées polaires r , ψ . Pour les coordonnées rectangulaires de ces vecteurs $x = (x_1, x_2)$ et $y = (y_1, y_2)$, on a alors

$$x_1 = \rho \cos \varphi, \quad x_2 = \rho \sin \varphi, \quad y_1 = r \cos \psi, \quad y_2 = r \sin \psi.$$

Donc, d'après la définition générale,

$$\cos \widehat{(x, y)} = \frac{(x, y)}{|x| \cdot |y|} = \frac{\rho r (\cos \varphi \cos \psi + \sin \varphi \sin \psi)}{\rho r} = \cos (\varphi - \psi).$$

Il en résulte que l'angle $\widehat{(x, y)}$ de deux vecteurs x et y est celui des nombres $\varphi - \psi + 2k\pi$ et $\psi - \varphi + 2k\pi$ qui est compris entre 0 et π .

5.75. **R o t a t i o n s.** a. Dans le plan R_2 , on appelle *rotation d'un angle* θ une transformation qui conserve l'origine des coordonnées et transforme tout vecteur $x \neq 0$ de coordonnées polaires ρ, φ en le vecteur x' de coordonnées polaires $\rho, \varphi + \theta$.

Exprimons la rotation en coordonnées rectangulaires. Soit, en coordonnées rectangulaires, $x = (x_1, x_2)$, $x' = (x'_1, x'_2)$. Nous

$$\begin{aligned}x_1' &= \rho \cos (\varphi + \theta) = \rho (\cos \varphi \cos \theta - \sin \varphi \sin \theta) = \\&= x_1 \cos \theta - x_2 \sin \theta, \\x_2' &= \rho \sin (\varphi + \theta) = \rho (\sin \varphi \cos \theta + \cos \varphi \sin \theta) = \\&= x_1 \sin \theta + x_2 \cos \theta;\end{aligned}$$

ce qui représente les formules cherchées de la rotation en coordonnées rectangulaires. Par définition une rotation conserve la longueur de tout vecteur. Montrons qu'une rotation conserve également l'angle de deux vecteurs. Si φ et ψ sont les angles polaires de deux vecteurs x et y , alors les angles polaires des vecteurs x' et y' obtenus par rotation sont $\varphi + \theta$ et $\psi + \theta$ respectivement. L'angle ω formé par les vecteurs x et y se détermine à partir de l'équation $\cos \omega = \cos(\varphi - \psi)$, l'angle ω' formé par les vecteurs x' et y' à partir de l'équation $\cos \omega' = \cos |(\varphi + \theta) - (\psi + \theta)|$; puisque ω et ω' sont les deux dans l'intervalle limité par 0 et π , alors $\omega = \omega'$, ce qu'il fallait démontrer. Le déterminant des coefficients de la transformation linéaire construite vaut, évidemment, 1.

b. Généralisons la définition d'une rotation au cas d'un espace euclidien n -dimensionnel. On appelle *rotation* ou *transformation orthogonale* une transformation linéaire $x' = u(x)$, ou bien, en coordonnées,

$$\left. \begin{aligned} x'_1 &= u_{11}x_1 + \dots + u_{1n}x_n, \\ &\vdots \\ x'_m &= u_{m1}x_1 + \dots + u_{mn}x_n. \end{aligned} \right\} \quad (1)$$

qui conserve les longueurs des vecteurs ainsi que les angles qu'ils forment, et pour laquelle $\det ||u_{j_k}|| = 1$.

Trouvons les conditions qu'il faut imposer aux coefficients $u_{j\lambda}$ pour que les propriétés formulées soient remplies. Soit $e_j = (0, \dots, 1, \dots, 0)$ ($j = 1, \dots, n$) et $e'_j = u(e_j)$. Les vecteurs $e'_1, \dots, e'_n$ doivent être orthogonaux et normés (2.68.a), de même que les vecteurs $e_1, \dots, e_n$. Il résulte de (1) que $e'_j = (u_{1j}, \dots, u_{nj})$, nous avons donc

$$(e_i, e_j) = \sum_{k=1}^n u_{ki} u_{kj} = \begin{cases} 1 & \text{pour } i=j, \\ 0 & \text{pour } i \neq j. \end{cases} \quad (2)$$

Les conditions (2) sont en même temps suffisantes pour que la transformation (1) soit une rotation. En effet, supposons que $x' = u(x)$,

$y' = u(y)$ soient déterminés d'après les formules (1); alors

$$\begin{aligned} (x', y') &= \sum_1^n x'_k y'_k = \sum_{k=1}^n \left(\sum_{i=1}^n u_{ki} x_i \right) \left(\sum_{j=1}^n u_{kj} y_j \right) = \\ &= \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{k=1}^n u_{ki} u_{kj} \right) x_i y_j = \sum_{i=1}^n x_i y_i = (x, y), \end{aligned}$$

d'où la conservation des longueurs des vecteurs et des angles qu'ils forment. On peut montrer que, pour $n = 2$, la définition donnée coïncide avec la définition a (exercice 19).

5.76. Coordonnées polaires dans l'espace. Calculons l'angle α_j formé par un vecteur $x = (x_1, \dots, x_n)$ et un vecteur de base $e_j = (0, \dots, 1, \dots, 0)$ dans un espace euclidien n -dimensionnel R_n . D'après la formule générale 5.74,

$$\cos \alpha_j = \cos (\widehat{x, e_j}) = \frac{(x, e_j)}{|x| \cdot |e_j|} = \frac{x_j}{|x|}.$$

Par conséquent, les coordonnées rectangulaires du vecteur x peuvent être mises sous la forme

$$x_j = |x| \cos (\widehat{x, e_j}), \quad x = |x| (\cos (\widehat{x, e_1}), \dots, \cos (\widehat{x, e_n})). \quad (i)$$

Les angles $\alpha_j = (\widehat{x, e_j})$ et $|x|$ sont appelés *coordonnées polaires du vecteur x* . Il existe la relation suivante entre les nombres $\alpha_1, \dots, \alpha_n$:

$$\sum_{j=1}^n \cos^2 \alpha_j = \sum \frac{x_j^2}{|x|^2} = 1.$$

L'angle des vecteurs x et y s'exprime par leurs angles polaires $\alpha_1, \dots, \alpha_n$ et $\beta_1, \dots, \beta_n$ suivant la formule

$$\cos (\widehat{x, y}) = \frac{(x, y)}{|x| |y|} = \cos \alpha_1 \cdot \cos \beta_1 + \dots + \cos \alpha_n \cdot \cos \beta_n.$$

Si $\omega = (\omega_1, \dots, \omega_n)$ est un vecteur unitaire, $|\omega| = 1$, alors les nombres ω_j , comme on le voit de (1), sont les cosinus des angles formés par le vecteur ω et chacun des vecteurs de base: $\omega_j = \cos (\widehat{\omega, e_j})$.

En particulier, les coefficients u_{jk} de la rotation 5.75 (1), en tant que coordonnées des vecteurs unitaires e_j , admettent une interprétation géométrique directe: un coefficient u_{ij} est le cosinus de l'angle que le vecteur e_j forme avec le vecteur de base e_i .

$x = x_0$, alors $\alpha(x)g(x)$ est une application de G dans R_m continue elle aussi pour $x = x_0$.

c. Si $y = f(x)$ est une application continue en un point $x = x_0$, alors $|f(x)|$ est une fonction numérique continue pour $x = x_0$.

5.84. Soit $w = f(z)$ une fonction définie sur un ensemble G dans le plan euclidien E_2 et dont le domaine de valeurs est aussi dans le plan E_2 . Alors $z = (x, y)$, $w = (u, v)$ peuvent être considérés comme nombres complexes. La fonction $w = f(z)$ s'appelle fonction complexe d'une variable complexe. Conformément à 5.81, la fonction $w = f(z)$ est continue en un point (non isolé) $z = a \in G$ si $\lim_{z \rightarrow a} f(z) = f(a)$, ou, autrement dit, si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $|z - a| < \delta$ implique

$$|f(z) - f(a)| < \varepsilon.$$

Conformément à 5.82, pour que la fonction $w = f(z) = u(z) + iv(z)$ soit continue pour $z = a$, il faut et il suffit que les deux fonctions à valeurs réelles $u(z)$ et $v(z)$ soient continues pour $z = a$.

D'après 5.83.a, b, si deux fonctions complexes $f(z)$ et $g(z)$ sont continues pour $z = a$, la somme $f(z) + g(z)$ et le produit de $f(z)$ par toute fonction réelle $\alpha(z)$ continue pour $z = a$ le sont aussi. En partant des opérations connues sur les nombres complexes, on peut ajouter à ces propriétés les suivantes :

a. Si $f(z)$ et $g(z)$ sont deux fonctions complexes continues pour $z = a$, le produit $f(z)g(z)$ l'est aussi pour $z = a$.

b. Si $f(z)$ est une fonction complexe continue pour $z = a$ et telle que $f(a) \neq 0$, alors $\frac{1}{f(z)}$ est continue elle aussi pour $z = a$.

c. Si $f(z)$ et $g(z)$ sont deux fonctions complexes continues pour $z = a$ et si $g(a) \neq 0$, alors $\frac{f(z)}{g(z)}$ est continue elle aussi pour $z = a$.

Ces propositions résultent des propriétés correspondantes des limites (4.73.e).

5.85. La fonction $w = z$ est, évidemment, continue pour tout $z = x + iy$. En vertu de 5.84, tout polynôme $w = P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n$ à coefficients complexes $a_0, \dots, a_n$ est une fonction continue de z , et toute fonction rationnelle

$$w = \frac{P(z)}{Q(z)} = \frac{a_0 z^n + \dots + a_n}{b_0 z^m + \dots + b_m}$$

est continue partout, sauf pour les valeurs de z telles que $Q(z) = 0$.

5.86. **Théorème sur l'existence d'une racine.** Un nombre z_0 qui satisfait à l'équation $P(z_0) = 0$ s'appelle *racine du polynôme* $P(z)$. Dans le domaine réel, il existe des polynômes qui ne possèdent aucune racine (exemple : $P(x) = x^2 + 1$). Le théorème

fondamental de l'algèbre affirme que, dans le domaine complexe, tout polynôme qui n'est pas constant possède au moins une racine. Nous allons démontrer ce théorème.

a. L e m m e. Soit un polynôme $P(z)$. Pour tout A positif, il existe un R tel que, pour $|z| \geq R$, on a l'inégalité $|P(z)| \geq A$.

D é m o n s t r a t i o n. Nous avons pour $a_0 \neq 0$ et $z \neq 0$:

$$\begin{aligned} P(z) &= a_0 z^n + a_1 z^{n-1} + \dots + a_n = \\ &= a_0 z^n \left[1 + \frac{a_1}{a_0 z} + \dots + \frac{a_n}{a_0 z^n} \right]. \end{aligned}$$

La limite de l'expression entre parenthèses, pour $z \rightarrow \infty$, valant 1 (4.73.f), il existe un nombre R_1 tel que, pour $|z| \geq R_1$,

$$\left| 1 + \frac{a_1}{a_0 z} + \dots + \frac{a_n}{a_0 z^n} \right| \geq \frac{1}{2}$$

et, par conséquent,

$$|P(z)| = |a_0| |z|^n \left| 1 + \frac{a_1}{a_0 z} + \dots + \frac{a_n}{a_0 z^n} \right| \geq \frac{1}{2} |a_0| |z|^n.$$

Si l'on pose maintenant $|z| \geq R > R_1$, alors l'inégalité

$$|P(z)| \geq \frac{1}{2} |a_0| |z|^n \geq \frac{1}{2} |a_0| R^n > A$$

est satisfaite pour $R > \sqrt[n]{2 \frac{A}{|a_0|}}$, ce qui démontre le lemme.

b. L e m m e. Soit $P(z)$ un polynôme de degré $n \geq 1$. Si $P(z_0) \neq 0$ et si un nombre $\delta_0 > 0$ est donné, alors il existe un point z_1 , $|z_1 - z_0| \leq \delta_0$ tel que $|P(z_1)| < |P(z_0)|$.

D é m o n s t r a t i o n. Considérons d'abord le cas où $z_0 = 0$, $P(0) = 1$. Soit a_k le premier des coefficients $a_{n-1}, \dots, a_0$ qui n'est pas nul. Alors

$$P(z) = 1 + a_k z^k + \dots + a_n z^n = 1 + a_k z^k [1 + H(z)],$$

où

$$H(z) = \frac{a_{k-1}}{a_k} z + \dots + \frac{a_0}{a_k} z^{n-k}.$$

Comme, évidemment, $H(0) = 0$, alors, en vertu de la continuité du polynôme $H(z)$, il existe un cercle $|z| \leq \delta \leq \delta_0$, $\delta^k |a_k| < 1$, dans lequel la condition

$$|H(z)| \leq \frac{1}{2} \quad (1)$$

est satisfaite. Trouvons z_1 de l'équation

$$z_1^k = -\delta^k \frac{|a_k|}{a_k}$$

qui est résoluble (5.73). Evidemment, $|z_1| = \delta$. Alors

$$\begin{aligned} |P(z_1)| &= |1 + a_h z_1^h + a_h z_1^h H(z_1)| = \\ &= |1 - \delta^h |a_h| - \delta^h |a_h| H(z_1)| \leq 1 - \delta^h |a_h| + \\ &+ \delta^h |a_h| |H(z_1)| \leq 1 - \delta^h |a_h| + \frac{1}{2} \delta^h |a_h| = \\ &= 1 - \frac{1}{2} \delta^h |a_h| < 1, \end{aligned}$$

ce qu'il fallait démontrer.

Dans le cas général, à l'aide de la relation

$$P(z) = \sum_{h=0}^n a_h z^{n-h} = \sum_{h=0}^n a_h [(z - z_0) + z_0]^{n-h},$$

nous disposons le polynôme $P(z)$ dans l'ordre des puissances décroissantes de $z - z_0$:

$$P(z) = \sum_{h=0}^n b_{n-h} (z - z_0)^{n-h}, \quad P(z_0) = b_0 \neq 0.$$

En y remplaçant $z - z_0$ par ζ , on obtient

$$P(z) = b_0 P_0(\zeta),$$

où

$$P_0(\zeta) = 1 + \frac{b_1}{b_0} \zeta + \dots + \frac{b_n}{b_0} \zeta^n.$$

Puisque $|P(z)| = |b_0| \cdot |P_0(\zeta)|$, le problème se ramène au précédent.

c. Nous sommes maintenant en mesure de démontrer le théorème fondamental.

T h é o r è m e. *Tout polynôme $P(z) = a_0 z^n + \dots + a_n$ ($n \geq 1$, $a_0 \neq 0$) possède, dans le plan complexe, au moins une racine.*

D é m o n s t r a t i o n. La fonction $|P(z)|$ est non négative, posons

$$\alpha = \inf_{z \in \mathbb{C}} \{|P(z)|\}.$$

Ensuite, en appliquant le lemme a on trouve un R tel que, pour $|z| \geq R$, on a l'inégalité

$$|P(z)| \geq \alpha + 1.$$

Donc, les valeurs de $P(z)$ à l'extérieur du cercle $|z| \leq R$ n'ont aucune importance pour le calcul de α . Nous pouvons noter

$$\alpha = \inf_{|z| \leq R} \{|P(z)|\}.$$

Or, le cercle $V = \{ |z| \leq R \}$, contrairement à tout le plan C , est déjà un ensemble compact (3.96.e). D'après le théorème 5.16.b, la fonction continue $|P(z)|$ atteint, en un point $z_0 \in V$, sa borne inférieure :

$$|P(z_0)| = \alpha.$$

Si $\alpha = 0$, alors $P(z_0) = 0$. Montrons que la supposition $P(z_0) \neq 0$ amène à une contradiction.

Le point z_0 appartient à l'intérieur du cercle V parce que, pour $|z| = R$, l'inégalité $|P(z)| \geq \alpha + 1$ est déjà valable. Ainsi, un certain petit cercle $|z - z_0| \leq \delta_0$ est entièrement inclus dans le cercle V . D'après le lemme b, on peut trouver dans ce cercle un point z_1 tel que $|P(z_1)| < |P(z_0)|$. Or, ceci veut dire que $\alpha = |P(z_0)|$ n'est pas la borne inférieure de la fonction $|P(z)|$ dans le cercle V . La contradiction obtenue montre que $P(z_0) = 0$, et le théorème est démontré.

5.87. On sait de l'algèbre [20] que, pour deux polynômes quelconques $P(z)$ et $Q(z) \neq 0$, on peut trouver, en divisant et en dégageant le reste, deux autres polynômes $S(z)$ et $R(z)$ tels que

$$P(z) = Q(z)S(z) + R(z) \quad (1)$$

de sorte que ou bien $R(z) = 0$, ou bien le degré de $R(z)$ (reste) est inférieur à celui du polynôme $Q(z)$; $P(z)$ est alors appelé *dividende*, $Q(z)$ *diviseur*, $S(z)$ *quotient*. La somme des degrés du diviseur et du quotient vaut le degré du dividende. Ici z est tout simplement un symbole, pas forcément un nombre, et le polynôme $P(z)$ est une expression formelle définie par l'ensemble des coefficients $a_0, \dots, a_n$, avec les règles naturelles d'addition et de multiplication. En substituant à z , dans l'égalité (1), un nombre complexe quelconque, cette égalité devient une égalité entre les nombres obtenus.

Soit, en particulier, $Q(z) = z - z_1$. Alors $R(z)$ est de degré nul, donc est une constante. Si, de plus, z_1 est une racine du polynôme $P(z)$, alors la constante $R(z)$ est nulle; en effet,

$$R(z) \equiv R(z_1) = P(z_1) - S(z_1)(z_1 - z_1) = 0.$$

La décomposition (1) prend la forme

$$P(z) = S(z)(z - z_1),$$

où le polynôme $S(z)$ a $n - 1$ pour degré.

En appliquant ce procédé au polynôme $S(z) = S_1(z)$, on obtient

$$S(z) = S_1(z) = S_2(z)(z - z_2),$$

z_2 étant une racine du polynôme $S_1(z)$, donc aussi du polynôme $P(z)$; le degré du polynôme $S_2(z)$ est $n - 2$. En continuant de la même manière, on peut abaisser les degrés des quotients successifs

jusqu'au zéro, et on aura la décomposition

$$P(z) = S_0 (z - z_1) \dots (z - z_n),$$

où S_0 est une constante et $z_1, \dots, z_n$ les racines du polynôme $P(z)$. Elles ne sont pas forcément distinctes deux à deux; en réunissant les parenthèses qui contiennent les mêmes racines, on aboutit à la forme définitive de la factorisation du polynôme

$$P(z) = S_0 (z - z_1)^{r_1} \dots (z - z_h)^{r_h}, \quad (2)$$

les nombres $z_1, \dots, z_h$ étant déjà tous différents. Les exposants $r_1, \dots, r_h$ sont appelés *multiplacités* des racines correspondantes. En multipliant les parenthèses, on peut voir que le nombre S_0 est égal au coefficient du terme du plus haut degré du polynôme $P(z)$.

5.88.a. Décomposition de fonctions rationnelles en éléments simples. Pour deux polynômes donnés $P_1(z)$ et $P_2(z)$ (à coefficients complexes, par exemple), on peut construire leur plus grand commun diviseur $D(z)$, c'est-à-dire un polynôme du plus haut degré possible, qui divise exactement chacun des polynômes $P_1(z)$ et $P_2(z)$ [20]. On démontre qu'il existe la représentation

$$D(z) = P_1(z) S_1(z) + P_2(z) S_2(z),$$

où $S_1(z)$ et $S_2(z)$ sont encore des polynômes en z .

Il est évident que toute racine du polynôme $D(z)$ en est une pour chacun des polynômes $P_1(z)$ et $P_2(z)$. Donc, si $P_1(z)$ et $P_2(z)$ n'ont aucune racine commune, leur plus grand commun diviseur est une constante qu'on peut poser égale à 1.

b. Considérons une fonction rationnelle

$$\frac{Q(z)}{P_1(z) P_2(z)},$$

$P_1(z)$ et $P_2(z)$ n'ayant aucune racine commune. D'après ce qui précède, il existe des polynômes $S_1(z)$ et $S_2(z)$ tels que

$$P_1(z) S_1(z) + P_2(z) S_2(z) \equiv 1,$$

et, par conséquent,

$$Q(z) = Q(z) P_1(z) S_1(z) + Q(z) P_2(z) S_2(z).$$

En divisant les deux membres de cette égalité par $P_1(z) P_2(z)$, on a

$$\frac{Q(z)}{P_1(z) P_2(z)} = \frac{Q(z) S_1(z)}{P_2(z)} + \frac{Q(z) S_2(z)}{P_1(z)}.$$

On peut continuer cette opération en décomposant $P_1(z)$ et $P_2(z)$ en produit de facteurs sans racines communes. Etant donnée la

représentation 5.87 (2) pour un polynôme $P(z)$, on peut construire ainsi la décomposition d'une fonction rationnelle $\frac{Q(z)}{P(z)}$ en somme

$$\frac{Q(z)}{P(z)} = \frac{Q(z)}{S_0(z-z_1)^{r_1} \dots (z-z_h)^{r_h}} = \sum_{j=1}^h \frac{Q_j(z)}{(z-z_j)^{r_j}}. \quad (1)$$

En disposant chaque numérateur $Q_j(z)$ dans l'ordre des puissances de la différence correspondante $z - z_j$, on peut passer ensuite de la décomposition (1) à

$$\frac{Q(z)}{P(z)} = T(z) + \sum_{j=1}^h \left\{ \sum_{m=1}^{r_j} \frac{A_{jm}}{(z-z_j)^m} \right\}, \quad (2)$$

où $T(z)$ est un polynôme et A_{jm} des nombres. La formule (2) donne la décomposition de la fonction rationnelle $\frac{Q(z)}{P(z)}$ en partie entière et éléments simples *).

5.89. Polynômes à coefficients réels.

a. Si $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0$ est un polynôme à coefficients réels et $\bar{z}_0$ le nombre complexe conjugué d'un nombre z_0 , alors, d'après 2.73,

$$\overline{P(z_0)} = \overline{a_n z_0^n + \dots + a_0} = a_n \bar{z}_0^n + \dots + a_0 = P(\bar{z}_0). \quad (1)$$

En particulier, si z_0 est une racine du polynôme $P(z)$, i.e. $P(z_0) = 0$, alors l'égalité (1) montre que $P(\bar{z}_0) = 0$ aussi, c'est-à-dire que $\bar{z}_0$ est lui aussi une racine du polynôme $P(z)$.

b. Soit $z_0 = x_0 + iy_0$, $\bar{z}_0 = x_0 - iy_0$; alors $(z - z_0)(z - \bar{z}_0) = (z - x_0 - iy_0)(z - x_0 + iy_0) = (z - x_0)^2 + y_0^2$. En divisant le polynôme $P(z)$ par le trinôme à coefficients réels $(z - x_0)^2 + y_0^2$, on obtient

$$P(z) = [(z - x_0)^2 + y_0^2] P_1(z),$$

où $P_1(z)$ est encore un polynôme à coefficients réels. Ceci permet de mettre la décomposition 5.87 (2) sous une forme réelle en réunissant les facteurs à racines complexes conjuguées; à la place de 5.87 (2), on a alors une décomposition en produit de facteurs à coefficients réels

$$P(z) = s_0 [(z - x_1)^2 + y_1^2]^{r_1} \dots [(z - x_l)^2 + y_l^2]^{r_l} \times \\ \times (z - x_{l+1})^{r_{l+1}} \dots (z - x_p)^{r_p}, \quad (2)$$

*) L'unicité d'une telle décomposition est aussi démontrée en algèbre; voir [20]. Le théorème d'unicité est à la base de la méthode des coefficients indéterminés qui permet de construire la décomposition (2). On peut trouver cette méthode dans [8].

où $x_1 \pm iy_1, \dots, x_l \pm iy_l$ sont les racines non réelles du polynôme $P(z)$, et $x_{l+1}, \dots, x_p$ ses racines réelles avec les multiplicités respectives $s_1, \dots, s_p$. En particulier, les racines non réelles $x_j + iy_j$ et $x_j - iy_j$ ont une même multiplicité.

e. D'une manière analogue, la décomposition 5.88 (2) en éléments simples pour deux polynômes $Q(z)$ et $P(z)$ à coefficients réels peut être mise sous la forme

$$\frac{Q(z)}{P(z)} = T(z) + \sum_{j=1}^l \left\{ \sum_{m=1}^{s_j} \frac{A_{jm} + zB_{jm}}{[(z-x_j)^2 + y_j^2]^m} \right\} + \sum_{j=l+1}^p \left\{ \sum_{m=1}^{s_j} \frac{A_{jm}}{(z-x_j)^m} \right\};$$

le lecteur n'a qu'à effectuer les transformations nécessaires

§ 5.9. Suites de fonctions

5.91. Considérons d'abord une suite des fonctions

$$f_1(x), f_2(x), \dots, f_n(x), \dots \quad (1)$$

données sur un ensemble quelconque E , à valeurs dans un espace métrique M . Nous dirons que la suite (1) *converge sur l'ensemble E* si toute suite de la forme

$$f_1(x_0), f_2(x_0), \dots, f_n(x_0), \dots \quad (x_0 \in E)$$

possède une limite dans M . Cette limite qui dépend, naturellement, du point x_0 sera désignée par $f(x_0)$; ainsi nous aurons une fonction $f(x)$ appelée *limite de la suite des fonctions $f_n(x)$* :

$$f(x) = \lim_{n \rightarrow \infty} f_n(x). \quad (2)$$

De la sorte, l'égalité (2) veut dire que, pour tout $\varepsilon > 0$ et tout point $x_0 \in E$, il existe un numéro N (qui dépend de ε et du point x_0) tel que, pour tous les $n \geq N$,

$$\rho[f(x_0), f_n(x_0)] \leq \varepsilon. \quad (3)$$

5.92. Exemples.

a. La suite de fonctions numériques

$$f_n(x) = \frac{1}{n} \cdot f(x)$$

converge vers zéro dans le domaine de définition de la fonction numérique $f(x)$.

b. La suite de fonctions numériques $f_n(x) = x^n$ converge dans l'intervalle $-1 < x \leq 1$. Sa limite $f(x)$ est une fonction qui vaut 0 pour $x < 1$ et 1 pour $x = 1$.

c. La suite de fonctions

$$f_n(x) = \begin{cases} n \sin nx & \text{pour } 0 \leq x \leq \frac{\pi}{n}, \\ 0 & \text{pour } \frac{\pi}{n} \leq x \leq \pi \end{cases}$$

converge en tout point $x \in [0, \pi]$ vers 0. Notons que, dans cet exemple, l'écart maximal de la fonction $f_n(x)$ de la fonction limite $f(x)$ qui vaut 0 ne converge pas vers 0 et, qui plus est, croît indéfiniment.

5.93. En général, la fonction limite d'une suite $f_n(x)$ peut ne pas posséder les propriétés propres aux fonctions $f_n(x)$. Ainsi, dans l'exemple 5.92.b, les fonctions $f_n(x)$ sont continues sur $[0, 1]$, tandis que la fonction limite est discontinue. Pour que l'on puisse dire quelque chose de définitif sur la fonction limite, il faut, habituellement, imposer quelques conditions au caractère de la convergence.

L'une des plus importantes conditions de ce genre est formulée dans la définition suivante.

D é f i n i t i o n. Une suite de fonctions 5.91 (1) est dite *uniformément convergente* vers sa fonction limite $f(x)$ sur E si, pour tout $\varepsilon > 0$, on peut trouver un numéro N tel que, quel que soit un point $x_0 \in E$, l'inégalité

$$\rho[f(x_0), f_n(x_0)] \leq \varepsilon \quad (1)$$

a lieu pour $n > N$.

Cette définition diffère de 5.91 : ici un même numéro N convient pour tous les points $x_0 \in E$, tandis que, dans la définition 5.91, le numéro N dépendait du choix du point x_0 . Donc, l'inégalité (1) doit être remplacée par l'inégalité

$$\sup_{x \in E} \rho[f(x), f_n(x)] \leq \varepsilon \quad (2)$$

pour tous les $n \geq N = N(\varepsilon)$.

Dans l'exemple 5.92.a, la convergence de la suite $f_n(x)$ vers sa limite est uniforme, et dans les exemples 5.92.b, c non.

L'utilité de la définition introduite est confirmée par les théorèmes qui suivent.

5.94. **T h é o r è m e.** Si chacune des fonctions $f_n(x)$ est bornée, c'est-à-dire que, pour un point $p \in M$, on a les inégalités

$$\rho[f_n(x), p] \leq C_n \quad (n = 1, 2, \dots),$$

et si la suite $f_n(x)$ converge uniformément sur E , alors la fonction limite $f(x)$ est aussi bornée.

D é m o n s t r a t i o n. Pour un $\varepsilon > 0$ donné, par exemple pour $\varepsilon = 1$, trouvons un numéro N de façon à avoir, pour $n \geq N$

et pour tout $x \in E$,

$$\rho |f_n(x), f(x)| \leq \varepsilon = 1.$$

D'où, pour tout $x \in E$,

$$\rho |f(x), \rho| \leq \rho |f(x), f_N(x)| + \rho |f_N(x), \rho| \leq 1 + C_N,$$

et nous voyons que la fonction $f(x)$ est bornée.

5.95. T h é o r è m e. *Supposons que E soit un espace métrique et que toute fonction de la suite $f_n(x)$ soit continue en un point $x_0 \in E$. Alors, si la suite $f_n(x)$ converge uniformément, la fonction limite $f(x)$ est continue elle aussi au point x_0 .*

D é m o n s t r a t i o n. Pour un $\varepsilon > 0$ donné, trouvons un numéro N de façon à avoir, pour tout $n \geq N$ et tout $x \in E$,

$$\rho |f_n(x), f(x)| < \frac{\varepsilon}{3}. \quad (1)$$

La fonction $f_N(x)$ étant continue pour $x = x_0$, il existe un $\delta > 0$ tel que, pour tout x de la boule $V = \{\rho(x, x_0) < \delta\}$, on a l'inégalité

$$\rho |f_N(x), f_N(x_0)| < \frac{\varepsilon}{3}. \quad (2)$$

Notons les inégalités (1) pour les points $x \in V$ et x_0 et pour $n = N$:

$$\rho |f_N(x_0), f(x_0)| < \frac{\varepsilon}{3}, \quad (3)$$

$$\rho |f_N(x), f(x)| < \frac{\varepsilon}{3}. \quad (4)$$

Il résulte de (2)-(4) que, pour $x \in V$,

$$\rho |f(x), f(x_0)| < \varepsilon,$$

ce qui exprime la continuité de la fonction $f(x)$ pour $x = x_0$.

5.96. En conséquence, nous avons :

T h é o r è m e. *La limite d'une suite uniformément convergente $f_n(x)$ de fonctions continues sur un espace métrique E est une fonction continue sur E .*

5.97. Lorsque l'espace métrique M est un espace réel m -dimensionnel R_m , de sorte que les fonctions $f_n(x)$ sont des fonctions à valeurs vectorielles, les inégalités 5.93 (1) et 5.93 (2) prennent respectivement la forme :

$$|f(x_0) - f_n(x_0)| \leq \varepsilon, \quad (1)$$

pour tout $x_0 \in E$ et pour $n \geq N(e, x_0)$;

$$\sup_{x \in E} |f(x_0) - f_n(x_0)| \leq e, \quad (2)$$

pour tout $n \geq N(e)$.

5.98. Dans le même cas ($M = R_m$), un critère de convergence uniforme analogue à celui de Cauchy 4.75 a lieu.

Critère de convergence uniforme de Cauchy. Pour qu'une suite des fonctions $f_n(x)$ définies sur un ensemble E et à valeurs dans un espace R_m converge uniformément sur l'ensemble E , il faut et il suffit que, pour tout $e > 0$, il existe un numéro N tel que, pour tous les $n \geq N$, $p \geq N$,

$$\sup_{x \in E} |f_n(x) - f_p(x)| \leq e.$$

Démonstration. Supposons que les fonctions $f_n(x)$ convergent uniformément sur E vers une fonction $f(x)$. Alors, pour un $e > 0$ donné, nous choisissons un numéro N tel que, pour tout $n \geq N$ et tout $x \in E$,

$$|f_n(x) - f(x)| \leq \frac{e}{2}.$$

En particulier, pour un $p \geq N$ et pour tout $x \in E$,

$$|f_p(x) - f(x)| \leq \frac{e}{2}.$$

Donc, pour tout $x \in E$ et pour n et p choisis,

$$|f_n(x) - f_p(x)| \leq e.$$

Par conséquent,

$$\sup_{x \in E} |f_n(x) - f_p(x)| \leq e.$$

Inversement, supposons que la condition de Cauchy soit remplie. En particulier, la suite des vecteurs $f_1(x_0), \dots, f_n(x_0), \dots$, avec un $x_0 \in E$ fixe, vérifie alors le critère de Cauchy dans R_m (4.75). Donc, en vertu du théorème 4.75, la suite $f_n(x_0)$ converge dans R_m vers un vecteur que nous désignons par $f(x_0)$. Avec tous les $x = x_0 \in E$ possibles, on obtient une fonction $f(x)$ à valeurs dans R_m . Montrons que $f_n(x)$ converge vers $f(x)$ uniformément.

Passons à la limite pour $p \rightarrow \infty$ dans l'inégalité

$$|f_n(x) - f_p(x)| \leq e.$$

La distance étant continue (5.12.b), nous avons, pour tous les $x \in E$ et $n \geq N$,

$$|f_n(x) - f(x)| < e,$$

ce qui exprime la convergence uniforme de $f_n(x)$ vers $f(x)$. Le théorème est démontré.

Remarque. Dans l'énoncé du théorème, on peut remplacer l'espace R_m par n'importe quel espace métrique complet. Nous reviendrons à ces questions dans le chapitre 12.

Exercices

1. Soit donnée une fonction numérique $y = f(x)$ définie dans un voisinage d'un point x_0 . On sait que, pour tout $\delta > 0$, il existe un $\varepsilon > 0$ tel que $|x - x_0| < \delta$ implique $|f(x) - f(x_0)| < \varepsilon$. Est-ce que la fonction $f(x)$ est continue pour $x = x_0$?

2. Supposons que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $|f(x) - f(x_0)| < \varepsilon$ implique $|x - x_0| < \delta$. Est-ce que la fonction $f(x)$ est continue pour $x = x_0$?

3. Démontrer la continuité de la fonction $y = |x|$.

4. Discuter la continuité des fonctions $y = (x)$ (partie fractionnaire de x) et $y = [x]$ (partie entière de x) (1.71).

5. Démontrer que la fonction de Riemann

$$f(x) = \begin{cases} \frac{1}{n} & \text{si } x = \frac{m}{n} \text{ est une fraction irréductible,} \\ 0 & \text{si } x \text{ est irrationnel,} \end{cases}$$

est continue en tout point irrationnel et discontinue en tout point rationnel de l'axe.

6. Si deux fonctions numériques $f(x)$ et $g(x)$ sont continues, alors les fonctions

$$\max\{f(x), g(x)\}, \min\{f(x), g(x)\}$$

le sont aussi.

7. Si une fonction numérique $f(x)$ est continue pour $x \in [a, b]$ et si $x_1, \dots, x_n$ sont des points arbitraires de cet intervalle, alors il existe un point $x_0 \in [a, b]$ tel que

$$f(x) = \frac{1}{n} [f(x_1) + \dots + f(x_n)].$$

8. Si une fonction monotone bornée $f(x)$ est continue sur un intervalle (a, b) , fini ou infini, alors elle est uniformément continue sur (a, b) .

9. Une fonction monotone $f(x)$ définie sur $(-\infty, \infty)$ et satisfaisant à l'équation fonctionnelle $f(x+y) = f(x) + f(y)$ est de la forme

$$f(x) = ax.$$

10. Indiquer, sur $(-\infty, \infty)$, les intervalles de convergence uniforme de la suite des fonctions

$$f_n(x) = \frac{1}{1+x^n}.$$

11. Démontrer que la relation 5.61 (4) n'est pas une conséquence des relations 5.61 (1)-(3), et la relation 5.61 (1) n'est pas une conséquence des relations 5.61 (2)-(4).

12. Une fonction $f(x)$ est définie pour $-\infty < x < \infty$, et, quels que soient deux points $x_1 < x_2$ et un $c \in (f(x_1), f(x_2))$, il existe un point $x_c \in (x_1, x_2)$ en lequel $f(x_c) = c$. Est-ce que la fonction $f(x)$ est continue?

13. Vérifier la table de valeurs :

x	$\sin x$	$\cos x$	$\operatorname{tg} x$
$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{3}$
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$

14. Soit une fonction réelle $f(x)$ définie sur un intervalle $[a, b]$. Démontrer que l'ensemble des points $c \in [a, b]$, pour lesquels $\lim_{x \rightarrow c} f(x)$ existe et est différent de $f(c)$, est au plus dénombrable.

15. Soit $t = 0, t_1 t_2 \dots t_n \dots$ la représentation décimale (1.77) d'un nombre réel t , $0 \leq t < 1$. Soit, ensuite, $n_1 < n_2 < \dots < n_k < \dots$ une sous-suite de la suite des nombres naturels. Discuter la continuité de la fonction

$$x(t) = 0, t_{n_1} t_{n_2} \dots t_{n_k} \dots$$

16. Démontrer que l'ensemble M des points du plan qui sont situés sur la circonférence unité Γ centrée à l'origine des coordonnées et qui ont les angles polaires $1, 2, \dots, n, \dots$ est partout dense sur Γ .

17. Pour qu'une fonction $f(x)$ définie et continue sur un ensemble P partout dense dans un compact M admette un prolongement continu sur tout M (i.e. pour qu'il existe une fonction $F(x)$ définie et continue sur M et identique à $f(x)$ sur P), il faut et il suffit que $f(x)$ soit uniformément continue sur P .

18. Démontrer qu'une fonction réelle $f(x)$ définie et non décroissante dans l'intervalle (a, b) peut avoir au plus un ensemble dénombrable de points de discontinuité.

19. Démontrer qu'une rotation dans l'espace n -dimensionnel (5.76.b) pour $n = 2$ est la rotation d'un certain angle dans le plan (5.76.a).

20. Si, dans la définition d'une rotation dans un espace n -dimensionnel (5.76.b), on omet la condition $\det ||u_{jk}|| = 1$, alors la transformation linéaire obtenue, si elle n'est pas rotation, est une « rotation avec réflexion », i.e. le résultat d'une réflexion ($\xi_1 = \xi_1, \dots, \xi_{n-1} = \xi_{n-1}, \xi_n = -\xi_n$) suivie d'une rotation.

Historique

La première définition correcte de la continuité pour les fonctions d'une variable réelle est donnée par Bolzano (1817), puis par Cauchy (1821). Les deux savants obtiennent le théorème sur la valeur intermédiaire (5.22) en appliquant le « critère de Cauchy ». Le théorème 5.16.b qui dit qu'une fonction continue est bornée et atteint ses bornes sur un compact est trouvé (pour un intervalle fermé) par Weierstrass aux environs de 1860. On doit à Heine (1870) la définition

de la continuité uniforme et le théorème correspondant sur les fonctions continues (5.17.b) (pour un intervalle fermé).

La fonction exponentielle d'exposant fractionnaire est connue depuis longtemps (Stifel, 1544) ; à cette époque lointaine, on remarque déjà qu'à un produit de puissances correspond la somme d'exposants, et que cette propriété peut être utilisée dans les calculs. Les premières tables de logarithmes sont dressées par Neper (1617) et Bürgli (1620). La définition complète de la fonction exponentielle et du logarithme ainsi que la démonstration de la continuité de ces fonctions sont données au milieu du XIX^e siècle, dès l'apparition des théories assez développées des nombres réels.

Les fonctions trigonométriques (en tant que segments dans un cercle) sont connues dès l'antiquité (la première table de sinus est établie par Ptolémée au II^e siècle). Les mathématiciens indiens connaissaient la formule 5.61 (1) (Aryabhata, V^e siècle). Les formules 5.61 (2) et (3) sont obtenues par le mathématicien allemand Regiomontanus (XV^e siècle). Euler (XVII^e siècle) introduit les notations modernes des fonctions trigonométriques.

Il existe une autre construction axiomatique de la théorie des fonctions trigonométriques qui ne suppose pas l'existence de solutions de certaines équations fonctionnelles et part des considérations de la théorie de groupes continus [16].

Le problème d'existence d'une racine complexe de tout polynôme à coefficients complexes se pose dès le XVII^e siècle ; au siècle suivant on en donne plusieurs solutions, dont aucune ne peut être rigoureuse faute d'une théorie développée de nombres réels et de continuité. La première démonstration rigoureuse qui précise le rôle de la continuité est donnée par Gauss (1799) ; les lacunes qui subsistent sont comblées par les travaux de Bolzano (1817), Cauchy (1821) et enfin de Dedekind-Cantor-Weierstrass (1874).

La notion de convergence uniforme d'une suite de fonctions et son rôle dans la conservation de la continuité sont établis par Stokes et Seidel (1847-1848), puis par Cauchy (1853) ; d'ailleurs un théorème analogue, pour un cas particulier, était déjà donné par Abel (1826).

Séries

En mettant de différentes manières des parenthèses dans l'expression $1 - 1 + 1 - 1 + \dots$, je peux, à mon gré, obtenir 0 ou 1. Mais alors l'idée de la création *ex nihilo* devient parfaitement plausible.

Guido Grandi (1703)

Quant à moi, j'avoue que tous les raisonnements et calculs basés sur des séries non convergentes... m'ont toujours l'air extrêmement suspect.

Jean d'Alembert (1748)

§ 6.1. Séries numériques. Séries à termes positifs

6.11.a. Soit $a_1, a_2, \dots, a_n, \dots$ une suite de nombres réels. Formons les *sommes partielles*

$$\begin{aligned} s_1 &= a_1, \\ s_2 &= a_1 + a_2, \\ &\dots \\ s_n &= a_1 + a_2 + \dots + a_n, \\ &\dots \end{aligned}$$

Définition. Si la suite des nombres $s_1, s_2, \dots, s_n, \dots$ converge vers une limite finie s , nous dirons que la *série numérique*

$$a_1 + a_2 + \dots + a_n + \dots \quad (1)$$

converge et que sa somme est le nombre $s = \lim_{n \rightarrow \infty} s_n$. Si la suite des nombres $s_1, s_2, \dots, s_n, \dots$ diverge, nous dirons que la *série* (1) *diverge* et nous ne lui associerons aucune somme.

Les nombres $a_1, a_2, \dots$ sont appelés *termes de la série* (1). Toute somme finie $a_{m+1} + \dots + a_n$ est un morceau de la série (1).

b. Exemple. Soit la série numérique

$$1 + x + x^2 + \dots + x^{n-1} + \dots, \quad (2)$$

où x est un nombre fixe. Les sommes partielles de cette série sont calculées d'après la formule de sommation d'une progression géométrique

$$s_n = \sum_{h=0}^{n-1} x^h = \frac{1-x^n}{1-x}.$$

Si $|x| < 1$, alors $x^n \rightarrow 0$ pour $n \rightarrow \infty$, d'où la convergence de s_n vers la limite

$$s = \frac{1}{1-x}. \quad (3)$$

Donc, pour $|x| < 1$, la série (2) converge et la valeur de sa somme est (3).

Si $x = 1$, alors, évidemment,

$$s_n = 1 + \dots + 1 = n,$$

la série (2) étant donc divergente. Pour $x = -1$, nous avons

$$s_1 = 1, \quad s_2 = 0, \quad s_3 = 1, \quad s_4 = 0, \dots,$$

de sorte que, tout en étant bornée, la suite s_n n'a pas de limite, et la série (2) diverge. Enfin, pour $|x| > 1$, la quantité $|s_n|$ croît indéfiniment avec n , de sorte que la série (2) diverge toujours.

En résumé: la série (2) converge pour $|x| < 1$ vers la somme (3) et diverge pour $|x| \geq 1$.

c. Critère de Cauchy pour une série. En appliquant le critère de Cauchy (3.72) à la suite des sommes partielles $s_1, s_2, \dots, s_n, \dots$ et compte tenu du fait que, pour $m < n$, on a $s_n - s_m = a_{m+1} + \dots + a_n$, nous obtenons le critère de Cauchy pour une série:

La série (1) converge si, et seulement si, pour tout $\varepsilon > 0$, il existe un nombre N tel que, quels que soient $m \geq N$, $n > m$,

$$|a_{m+1} + \dots + a_n| < \varepsilon.$$

d. En particulier, lorsque la série (1) converge, alors, pour tout $\varepsilon > 0$, il existe un N tel que

$$|a_n| < \varepsilon,$$

dès que $n \geq N$, c'est-à-dire que $\lim_{n \rightarrow \infty} a_n = 0$.

Donc, pour une série convergente $a_1 + \dots + a_n + \dots$, la suite de ses termes $a_1, \dots, a_n, \dots$ tend vers zéro. La relation $a_n \rightarrow 0$ représente donc une condition nécessaire de convergence de la série $a_1 + \dots + a_n + \dots$. Il existe des séries $a_1 + \dots + a_n + \dots$ divergentes, pour lesquelles $a_n \rightarrow 0$ (exemple dans 6.15.b), de sorte que la condition $a_n \rightarrow 0$ n'est pas une condition suffisante de convergence d'une série.

e. Une autre condition nécessaire de convergence d'une série découle de 3.33.b: toute suite convergente est bornée. Donc, pour

toute série convergente $a_1 + \dots + a_n + \dots$, la suite de ses sommes partielles est bornée. Tout en étant une condition nécessaire de convergence, cette condition n'est pas une condition suffisante. Ainsi, les sommes partielles de la série $1 - 1 + 1 - 1 + \dots$ sont bornées, sans qu'elle soit convergente.

6.12. Aux numéros suivants (6.13-6.17), nous considérons les séries à termes non négatifs. Pour une telle série, les sommes $s_1, s_2, \dots$ forment une suite non décroissante. Donc, si les sommes partielles $s_1, s_2, \dots$ d'une série à termes non négatifs sont bornées (pour $n \rightarrow \infty$), alors la série est convergente. Nous voyons que, pour une série à termes non négatifs, la condition nécessaire 6.11.e est aussi suffisante.

6.13.a. Cette propriété est à la base du critère de comparaison: Si une série à termes non négatifs $a_1 + a_2 + \dots + a_n + \dots$ converge, alors il en est de même de toute série à termes non négatifs

$$b_1 + b_2 + \dots + b_n + \dots,$$

s'il existe une constante c et un numéro N tels que, pour $n \geq N$, on a les inégalités $b_n \leq ca_n$.

En effet, soient $s = a_1 + a_2 + \dots$ et $\sigma_n = b_1 + \dots + b_n$; nous avons, pour $n > N$:

$$\begin{aligned} \sigma_n &= b_1 + \dots + b_n \leq b_1 + \dots + b_N + \\ &+ c(a_{N+1} + \dots + a_n) \leq b_1 + \dots + b_N + cs. \end{aligned}$$

b. Exemple. Si les termes d'une série à termes non négatifs $b_1 + b_2 + \dots$ vérifient les inégalités

$$b_n \leq c\theta^n \quad (n = N, N+1, \dots),$$

où $c > 0$ et $0 \leq \theta < 1$, alors la série $b_1 + b_2 + \dots$ converge.

En effet, on peut appliquer le critère de comparaison et 6.11.b en posant $a_n = \theta^n$.

6.14. Les deux critères qui suivent sont basés sur les résultats de 6.13.

a. Critère de d'Alembert. Si une série à termes positifs

$$a_1 + a_2 + \dots + a_n + a_{n+1} + \dots \quad (1)$$

vérifie l'inégalité

$$\overline{\lim}_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} < 1, \quad (2)$$

alors la série (1) converge. Si

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} > 1, \quad (3)$$

la série (1) diverge.

Démonstration. Etant donné (2), on a, pour un $\theta < 1$ et à partir d'un numéro N ,

$$\frac{a_{n+1}}{a_n} \leq \theta \quad (n \geq N).$$

Donc

$$a_{N+1} \leq \theta a_N, \quad a_{N+2} \leq \theta a_{N+1} \leq \theta^2 a_N, \dots, \quad a_{N+k} \leq \theta^k a_N, \dots,$$

et la convergence de la série (1) découle de 6.13. Etant donné (3), on a, à partir d'un numéro N ,

$$\frac{a_{n+1}}{a_n} > 1 \quad (n = N, N+1, \dots);$$

donc $a_{N+1} > a_N$, $a_{N+2} > a_{N+1}$, $\dots$, de sorte que les termes de la série ne tendent pas vers 0; par conséquent, la série (1) diverge en vertu de 6.11.d.

Exemple. La série

$$1 + \frac{1 \cdot 2}{1 \cdot 3} + \frac{1 \cdot 2 \cdot 3}{1 \cdot 3 \cdot 5} + \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 3 \cdot 5 \cdot 7} + \dots$$

converge parce que

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{2n+1} \rightarrow \frac{1}{2}.$$

b. Critère de Cauchy. Si une série à termes positifs

$$a_1 + a_2 + \dots + a_n + \dots \quad (4)$$

vérifie l'inégalité

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} < 1, \quad (5)$$

alors la série (4) converge. Si

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} > 1, \quad (6)$$

la série (4) diverge.

Démonstration. Une fois (5) remplie, l'inégalité

$$\sqrt[n]{a_n} < \theta$$

a lieu pour un $\theta < 1$ et à partir d'un numéro N , d'où, pour $n \geq N$,

$$a_n < \theta^n,$$

et la série (4) converge d'après 6.13. Une fois (6) remplie, l'inégalité

$$\sqrt[n]{a_n} > 1$$

a lieu pour un $n > N$, quel que soit N . Il en résulte que

$$a_n > 1;$$

donc, les termes de la série (4) ne tendent pas vers 0, et la série diverge d'après 6.11.d.

Exemple. La série

$$\frac{2}{1} + \left(\frac{3}{3}\right)^2 + \left(\frac{4}{5}\right)^3 + \left(\frac{5}{7}\right)^4 + \dots$$

converge puisque

$$\sqrt[n]{a_n} = \frac{n+1}{2n-1} \rightarrow \frac{1}{2}.$$

6.15. Signalons encore un critère de convergence utile dû à Cauchy.

Si $a_1 \geq a_2 \geq \dots \geq a_n \geq \dots \geq 0$, la série $\sum_1^\infty a_n$ converge ou diverge en même temps que la série

$$\sum_1^\infty 2^h a_{2^h}. \quad (1)$$

En effet, considérons la somme $\sum_1^m a_j$, et soit un nombre k tel que $2^k > m$. Alors

$$\begin{aligned} a_1 + \dots + a_m &\leq a_1 + \dots + a_{2^k-1} = a_1 + (a_2 + a_3) + \dots + \\ &+ (a_{2^{k-1}-1} + \dots + a_{2^k-1}) \leq a_1 + 2a_2 + \dots + 2^{k-1}a_{2^{k-1}}. \end{aligned}$$

Si la série $\sum_1^\infty 2^h a_{2^h}$ converge, le second membre ne dépasse pas sa somme. Donc, les sommes partielles $a_1 + \dots + a_m$ sont bornées, et $\sum_1^\infty a_n$ converge. Au contraire, si k est tel que $2^k < m$, nous avons

$$\begin{aligned} a_1 + \dots + a_m &\geq a_1 + \dots + a_{2^k} = a_1 + a_2 + (a_3 + a_4) + \dots + \\ &+ (a_{2^{k-1}-1} + \dots + a_{2^k}) \geq \frac{1}{2} a_1 + a_2 + 2a_4 + \dots + 2^{k-1}a_{2^k} = \\ &= \frac{1}{2} (a_1 + 2a_2 + 4a_4 + \dots + 2^k a_{2^k}), \end{aligned}$$

et la divergence de la série $\sum_1^\infty 2^h a_{2^h}$ implique celle de la série $\sum_1^\infty a_n$.

Exemples. a. La série $\sum_1^\infty \frac{1}{n^p}$ converge pour $p > 1$ et diverge pour $0 \leq p \leq 1$.

En effet, si $p \geq 0$, les termes de la série ne croissent pas, et on peut appliquer le critère de Cauchy. La série (1) est alors de

la forme

$$\sum_1^{\infty} \frac{2^n}{2^{np}} = \sum_1^{\infty} 2^{n(1-p)} = \sum_1^{\infty} \alpha^n, \text{ où } \alpha = 2^{1-p},$$

et le résultat cherché découle de l'exemple 6.13.b, vu que, pour $p > 1$, on a $2^{1-p} < 1$, et, pour $p \leq 1$, on a $2^{1-p} \geq 1$.

Notons que

$$\frac{1}{(n+1)^p} : \frac{1}{n^p} = \frac{1}{\left(\frac{n+1}{n}\right)^p} = \frac{1}{\left(1 + \frac{1}{n}\right)^p} \rightarrow 1 \quad (n \rightarrow \infty),$$

de sorte que le critère de d'Alembert n'est pas applicable au cas considéré.

b. En particulier, la série $\sum_1^{\infty} \frac{1}{n}$ *diverge*. Chaque terme de cette série, à partir du deuxième, est la moyenne harmonique de deux termes voisins*), c'est la raison pour laquelle la série $\sum_1^{\infty} \frac{1}{n}$ est dite *harmonique*.

c. La série $\sum_{n=2}^{\infty} \frac{1}{n \log_a^n n}$ *converge pour $p > 1$ et diverge pour $0 \leq p \leq 1$ (si $a > 1$)*. Dans ce cas aussi, vu la non-décroissance du logarithme (5.42), on peut appliquer le même critère de Cauchy. La série (1) a alors la forme

$$\sum \frac{2^n}{2^n (\log_a 2^n)^p} = \sum \frac{1}{(n \log_a 2)^p} = \frac{1}{(\log_a 2)^p} \sum \frac{1}{n^p},$$

et on est ramené à l'exemple a.

6.16. La convergence ou la divergence de nombreuses séries, pour lesquelles les critères de d'Alembert 6.14.a et de Cauchy 6.14.b ne jouent pas, peut être établie en comparant une telle série avec une série 6.15.a ou c.

Remarquons au préalable que ce ne sont pas seulement les termes de séries qu'on peut comparer, mais aussi les rapports de termes voisins.

*) Un nombre a s'appelle *moyenne harmonique* de deux nombres b et c si $\frac{1}{a} = \frac{1}{2} \left(\frac{1}{b} + \frac{1}{c} \right)$.

L e m m e. *Etant données deux séries à termes positifs $\sum_1^{\infty} u_n$ et $\sum_1^{\infty} v_n$, si l'inégalité*

$$\frac{u_{n+1}}{u_n} \leq \frac{v_{n+1}}{v_n} \quad (1)$$

a lieu pour tous les n assez grands, et si la série $\sum_1^{\infty} v_n$ converge, alors il en est de même de la série $\sum_1^{\infty} u_n$.

Démonstration. Supposons que l'inégalité (1) soit vérifiée par $n = N, N+1, N+2, \dots$. En multipliant membre à membre les inégalités (1) écrites pour $n = N, N+1, \dots, N+p$, on trouve

$$\frac{u_{N+p}}{u_N} \leq \frac{v_{N+p}}{v_N},$$

ou bien

$$u_{N+p} \leq \frac{u_N}{v_N} v_{N+p}; \quad (2)$$

l'inégalité (2) étant valable pour tous les $p = 1, 2, \dots$, la convergence de la série $\sum_1^{\infty} u_n$ résulte de 6.13.

6.17.a. Théorème (critère de Raabe). *Si, pour une série à termes positifs $\sum_1^{\infty} u_n$, le rapport de deux termes voisins peut être mis sous la forme*

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha E}{n} \quad (E \rightarrow 1 \text{ pour } n \rightarrow \infty, \text{ cf. 4.39}),$$

alors la série $\sum_1^{\infty} u_n$ converge pour $\alpha > 1$ et diverge pour $\alpha < 1$.

Démonstration. Posons $v_n = \frac{1}{n^\alpha}$; alors, d'après 5.59.f, g,

$$\frac{v_{n+1}}{v_n} = \frac{n^\alpha}{(n+1)^\alpha} = \frac{1}{\left(1 + \frac{1}{n}\right)^\alpha} = \frac{1}{1 + \frac{\alpha E}{n}} = 1 - \frac{\alpha E}{n}.$$

Si $\alpha > 1$, choisissons un $p \in (1, \alpha)$; pour les n assez grands, nous avons

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha E}{n} < 1 - \frac{pE}{n} = \frac{v_{n+1}}{v_n}.$$

La série $\sum_1^{\infty} v_n$ étant convergente (6.15.a), la convergence de la série $\sum_1^{\infty} u_n$ résulte de 6.16. Si $\alpha < 1$, choisissons un $p \in (\alpha, 1)$; pour les n assez grands, nous avons

$$\frac{v_{n+1}}{v_n} = 1 - \frac{pE}{n} < 1 - \frac{\alpha E}{n} = \frac{u_{n+1}}{u_n}.$$

Si la série $\sum_1^{\infty} u_n$ convergerait, la série $\sum_1^{\infty} v_n$ elle aussi convergerait d'après 6.16. Or, cette dernière diverge ($p < 1$), donc, la série $\sum_1^{\infty} u_n$ diverge.

b. Exemple. Considérons la série (γ et δ sont des nombres autres que 0, -1 , -2 , ..., $-n$, ...)

$$\sum_{n=1}^{\infty} \frac{\gamma(\gamma+1) \dots (\gamma+n-1)}{\delta(\delta+1) \dots (\delta+n-1)}. \quad (1)$$

Dans ce cas, $\frac{u_{n+1}}{u_n} = \frac{\gamma+n}{\delta+n} = \frac{1+\frac{\gamma}{n}}{1+\frac{\delta}{n}} = 1 - \frac{\delta-\gamma}{n} E$. D'après a, la série (1) converge si $\delta-\gamma > 1$ et diverge si $\delta-\gamma < 1$.

§ 6.2. Séries à termes réels de signes quelconques

6.21. Pour une série

$$a_1 + a_2 + \dots + a_n + \dots \quad (1)$$

à termes réels quelconques, considérons la série à termes non négatifs

$$|a_1| + |a_2| + \dots + |a_n| + \dots \quad (2)$$

Théorème. Si la série (2) converge, la série (1) converge elle aussi.

Démonstration. Pour un $\varepsilon > 0$ donné, trouvons un numéro N de façon à avoir

$$|a_{m+1}| + \dots + |a_n| < \varepsilon$$

pour $n > m \geq N$. Les m et n ainsi choisis, nous avons

$$|a_{m+1} + \dots + a_n| \leq |a_{m+1}| + \dots + |a_n| < \varepsilon,$$

de sorte que la condition de Cauchy 6.11.c est remplie pour la série (1) ce qui démontre le théorème.

6.22. Une série (1), pour laquelle la série (2) converge, est dite *absolument convergente*.

Il peut arriver qu'une série (1) converge, tandis que la série (2) diverge (on en donnera un exemple dans la suite); une telle série (1) est dite *semi-convergente*.

6.23. Critère de Leibniz. Si $a_1 \geq a_2 \geq a_3 \geq \dots$ et $a_n \rightarrow 0$, alors la série alternée

$$a_1 - a_2 + a_3 - a_4 + \dots$$

converge.

Démonstration. Nous avons

$$s_{2n+1} = s_{2n-1} - a_{2n} + a_{2n+1} \leq s_{2n-1},$$

$$s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geq s_{2n},$$

de sorte que la suite $s_2, s_4, \dots$ ne décroît pas et la suite $s_1, s_3, \dots$ ne croît pas. Pour tout k et pour $n \geq k$,

$$s_2 \leq s_4 \leq \dots \leq s_{2n} \leq s_{2n} + a_{2n+1} = s_{2n+1} \leq \dots \leq s_{2k+1},$$

de sorte que la suite $s_2, s_4, \dots$ est majorée par n'importe quel s_{2k+1} ; par conséquent, $\xi = \lim_{n \rightarrow \infty} s_{2n} \leq s_{2k+1}$ pour tout k . Ceci

veut dire que $\lim_{n \rightarrow \infty} s_{2n}$ sert de minorant à tous les s_{2k+1} , il existe donc $\lim_{n \rightarrow \infty} s_{2n+1}$, et

$$\xi = \lim_{n \rightarrow \infty} s_{2n} \leq \lim_{n \rightarrow \infty} s_{2n+1} = \eta.$$

Ensuite, nous avons $0 \leq \eta - \xi \leq s_{2n+1} - s_{2n} = a_{2n+1}$; comme, par hypothèse, $a_{2n+1} \rightarrow 0$, alors $\xi = \eta = \lim_{n \rightarrow \infty} s_n$ (4.64), ce qu'il fallait démontrer.

6.24. Exemples.

a. La série

$$1 - \frac{1}{2^\alpha} + \frac{1}{3^\alpha} - \frac{1}{4^\alpha} + \dots \quad (\alpha > 0)$$

converge d'après le critère de Leibniz. Lorsque $\alpha > 1$, elle converge absolument (6.15.a). Lorsque $0 < \alpha \leq 1$, elle n'est que semi-convergente, puisque la série des valeurs absolues

$$1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \frac{1}{4^\alpha} + \dots$$

diverge (6.15.a).

b. En particulier, la série alternée

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

converge. Dans ce qui suit (9.74), nous donnerons la valeur de sa somme.

c. Soit la série

$$\sum_{n=1}^{\infty} (-1)^n u_n, \quad u_n = \frac{\gamma(\gamma+1)\dots(\gamma+n-1)}{\delta(\delta+1)\dots(\delta+n-1)},$$

$$\gamma, \delta \neq 0, -1, -2, \dots$$

Cette série converge absolument pour $\delta > \gamma + 1$ (6.17.b). Montrons qu'elle converge pour $\delta > \gamma$ (au sens de la semi-convergence). Comme

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\delta - \gamma}{n} E'$$

(6.17.b), si $\delta > \gamma$, nous avons $u_{n+1} < u_n$ au moins pour les n assez grands. Pour appliquer le critère de Leibniz, il reste à établir que $\lim_{n \rightarrow \infty} u_n = 0$. Vu 5.59.d et g, nous avons

$$\ln \frac{u_{n+1}}{u_n} = \ln \left(1 - \frac{\delta - \gamma}{n} E' \right) = -\frac{\delta - \gamma}{n} E' \quad (E' \rightarrow 1 \text{ pour } n \rightarrow \infty);$$

$$\ln \frac{u_{n+1}}{u_1} = \ln \frac{u_2}{u_1} + \dots + \ln \frac{u_{n+1}}{u_n} = -(\delta - \gamma) \sum_{k=1}^n \frac{E'_k}{k},$$

d'où, étant donné la divergence de la série harmonique (6.15.b), $\ln u_{n+1} \rightarrow -\infty$, donc, $u_{n+1} \rightarrow 0$, ce qu'il fallait démontrer.

§ 6.3. Opérations sur les séries

6.31. Définitions. Soient deux séries numériques

$$a_1 + a_2 + \dots + a_n + \dots, \quad (1)$$

$$b_1 + b_2 + \dots + b_n + \dots \quad (2)$$

a. La série

$$(a_1 + b_1) + (a_2 + b_2) + \dots + (a_n + b_n) + \dots \quad (3)$$

s'appelle *somme des séries (1) et (2)*.

b. Si α est un nombre, alors la série

$$\alpha a_1 + \alpha a_2 + \dots + \alpha a_n + \dots \quad (4)$$

s'appelle *produit de la série (1) par le nombre α* .

c. La série

$$a_1 b_1 + (a_1 b_2 + b_1 a_2) + (a_1 b_3 + a_2 b_2 + a_3 b_1) + \dots \quad (5)$$

s'appelle *produit des séries (1) et (2)*.

Toutes ces définitions sont bien formelles si l'on n'introduit aucune hypothèse sur la convergence. C'est dans le théorème suivant qu'elles acquièrent un sens intuitif.

6.32. Théorème.

a. Si les séries (1) et (2) convergent et $\sum_1^{\infty} a_k = A$, $\sum_1^{\infty} b_k = B$, alors la série (3) converge elle aussi, et $\sum_1^{\infty} (a_k + b_k) = A + B$.

b. Si la série (1) converge et $\sum_1^{\infty} a_k = A$, alors la série (4) converge elle aussi et $\sum_1^{\infty} \alpha a_k = \alpha A$.

c. Si les séries (1) et (2) convergent, $\sum_1^{\infty} a_k = A$, $\sum_1^{\infty} b_k = B$ et l'une au moins de ces séries converge absolument, alors la série (5) converge et sa somme vaut $A \cdot B$.

Démonstration. (a) Si $A_n = \sum_1^n a_k$, $B_n = \sum_1^n b_k$, alors $\sum_1^n (a_k + b_k) = A_n + B_n \rightarrow A + B$ pour limite.

(b) Si $A_n = \sum_1^n a_k$, alors $\sum_1^n \alpha a_k = \alpha A_n \rightarrow \alpha A$ pour limite.

(c) Posons $A_n = \sum_1^n a_k$, $B_n = \sum_1^n b_k$, $c_n = \sum_1^n a_k b_{n-k+1}$, $C_n = \sum_1^n c_k$. Supposons que la série $\sum_1^{\infty} b_k$ converge absolument, de sorte que $D = \sum_1^{\infty} |b_k| < \infty$. Les nombres A_n sont bornés car la suite $A_1, A_2, \dots$ converge; soit $M = \sup |A_n|$. Les nombres $|A_n - A_m| \leq |A_n| + |A_m|$ sont bornés par le nombre $2M$. Pour un $\varepsilon > 0$ donné, trouvons un numéro N tel que, pour $n > m \geq N$

$\left| \sum_{k=m+1}^n a_k \right| \leq \frac{\varepsilon}{2D}$ et $\sum_{k=N}^{\infty} |b_k| \leq \frac{\varepsilon}{4M}$. Alors, pour $n \geq 2N$,

$|A_n B_n - C_n| = |(a_2 + \dots + a_n) b_n + (a_3 + \dots + a_n) b_{n-1} + \dots + a_n b_2| \leq |a_2 + \dots + a_n| \cdot |b_n| + |a_3 + \dots + a_n| \cdot |b_{n-1}| + \dots + |a_n| \cdot |b_2| \leq 2M(|b_n| + \dots + |b_{N+1}|) + \frac{\varepsilon}{2D}(|b_2| + \dots + |b_N|) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$

de sorte que $\lim_{n \rightarrow \infty} C_n = \lim_{n \rightarrow \infty} A_n B_n = A \cdot B$, et le théorème est démontré.

b. Théorème. Si $a_n \rightarrow 0$ et si le nombre d'éléments de chaque groupe est borné par une même constante (i.e. $m_n - m_{n-1} \leq M$, pour tout n), alors la convergence de la série (2) implique celle de la série (1).

En effet, pour un n donné, trouvons m_k et m_{k+1} de façon à avoir

$$m_k \leq n \leq m_{k+1}.$$

Alors, si $s_n = a_1 + \dots + a_n$, $\sigma_k = a_1 + \dots + a_{m_k}$, nous avons

$$|s_n - \sigma_k| = |a_{m_k+1} + \dots + a_n| \leq M \cdot \max_{j > m_k} |a_j| \rightarrow 0,$$

pour $k \rightarrow \infty$; il en résulte que les nombres s_n ont la même limite que σ_k .

6.35. Permutations de termes d'une série. Soit une série

$$a_1 + a_2 + \dots + a_n + \dots \quad (1)$$

Définition. Si $m_1, m_2, \dots$ est une permutation de l'ensemble des nombres naturels, alors la série

$$b_1 + b_2 + \dots + b_n + \dots, \quad (2)$$

où $b_n = a_{m_n}$ ($n = 1, 2, \dots$), est appelée *permutation de la série* (1).

Théorème. Si la série (1) converge et $a_n \geq 0$ ($n = 1, 2, \dots$), alors la série (2) converge vers la même somme.

Démonstration. Pour toute somme partielle

$$\sigma_n = b_1 + \dots + b_n \quad (3)$$

de la série (2), il existe une somme partielle

$$s_{m(n)} = a_1 + a_2 + \dots + a_{m(n)} \quad (4)$$

de la série (1) comprenant tous les termes de la somme partielle (3). De même, il existe une somme partielle

$$\sigma_{N(n)} = b_1 + \dots + b_n + \dots + b_{N(n)}$$

de la série (2) qui comprend tous les termes de la somme partielle (4). Comme $a_n \geq 0$ pour tout n , nous avons

$$\sigma_n \leq s_{m(n)} \leq \sigma_{N(n)}. \quad (5)$$

La première de ces inégalités a pour conséquence

$$\sigma_n \leq s_{m(n)} \leq s,$$

de sorte que la série (2) converge; désignons sa somme par σ . Il est évident que $m(n)$ et $N(n)$ sont au moins égaux à n , donc croissent indéfiniment avec n . En passant à la limite dans (5), pour

$n \rightarrow \infty$, nous obtenons

$$\sigma \leq s \leq \sigma,$$

d'où $\sigma = s$, le théorème étant ainsi démontré.

6.36. Théorème (Dirichlet). *Si la série 6.35 (1) converge absolument, la série 6.35 (2) converge absolument vers la même somme.*

Démonstration. La convergence absolue de la série (2) résulte de 6.35. Pour un $\varepsilon > 0$ donné, trouvons un numéro N tel

que $|s_N - s| < \frac{\varepsilon}{2}$ et que l'inégalité $\sum_{m=N}^n |a_m| \leq \frac{\varepsilon}{2}$ ait lieu pour

$m \geq N$. Soit p_0 un nombre tel que toute somme partielle σ_p de la série (2), avec $p \geq p_0$, comprenne les N premiers termes de la série (1). Alors, pour tout $p \geq p_0$, la différence $\sigma_p - s_N$ ne comprend que les termes de la série (1) de numéros supérieurs à N ; donc

$$|s_N - \sigma_p| \leq \frac{\varepsilon}{2}, \text{ d'où}$$

$$|s - \sigma_p| \leq |s - s_N| + |s_N - \sigma_p| \leq \varepsilon.$$

Ceci veut dire que $\sigma_p \rightarrow s$, ce qu'il fallait démontrer.

6.37. Le cas des séries semi-convergentes est tout différent.

Théorème (Riemann). *Etant donnée une série 6.35 (1) semi-convergente et deux nombres quelconques $\alpha \leq \beta$ (sur la droite réelle achevée), il existe une permutation 6.35 (2) de la série 6.35 (1) telle que*

$$\lim \sigma_n = \alpha, \quad \overline{\lim} \sigma_n = \beta.$$

Démonstration. La série 6.35 (1) étant convergente, ses termes tendent vers 0. Il n'existe donc qu'un nombre fini des numéros n pour lesquels $|a_n|$ dépasse un nombre positif donné ε . Ceci permet de choisir, dans tout ensemble de termes de la série 6.35 (1), le plus grand (ou le plus grand en module). En séparant, dans cette série, les termes positifs et les termes négatifs et en choisissant successivement les termes les plus grands (ou les plus grands en module), nous sommes en mesure de former deux séries: l'une

$$p_1 + p_2 + \dots + p_n + \dots \quad (1)$$

composée de tous les termes positifs de la série 6.35 (1) rangés dans l'ordre de décroissance, de sorte que $p_1 \geq p_2 \geq \dots$; et l'autre

$$q_1 + q_2 + \dots + q_n + \dots \quad (2)$$

composée des valeurs absolues de tous les termes négatifs de la série 6.35 (1) rangés eux aussi dans l'ordre de décroissance, de sorte que $q_1 \geq q_2 \geq \dots$.

Les deux séries sont divergentes.

En effet, si les deux séries (1) et (2) convergent vers les sommes respectives s et σ , alors la série 6.35 (1) convergerait absolument, puisque toute somme finie $|a_1| + \dots + |a_n|$ ne dépasserait pas la somme $s + \sigma$.

Abordons le cas où l'une des séries (1), (2) converge et l'autre non. Supposons, par exemple, que la série (2) converge et la série (1) diverge. Les sommes partielles s_n de la série 6.35 (1) comprennent, lorsque n croît, un nombre de plus en plus grand de termes des séries (1) et (2); les termes de la série (1) donnent une somme arbitrairement grande, tandis que ceux de la série (2) ne peuvent fournir qu'une somme limitée. Ainsi, $s_n \rightarrow \infty$ et la série 6.35 (1) ne peut pas être convergente. Une situation analogue à condition de changer de signe, a lieu si la série (1) converge et la série (2) diverge. Donc, les deux séries (1) et (2) divergent.

A présent, nous construisons avec les termes des séries (1) et (2) une nouvelle série d'après la règle suivante (pour le cas où α et β sont finis). Les sommes partielles s_n^* de la série (1) croissent indéfiniment; trouvons un numéro n_1 de façon à avoir

$$p_1 + \dots + p_{n_1-1} < \beta < p_1 + \dots + p_{n_1}$$

(si l'on a déjà $p_1 > \beta$, on pose $n_1 = 1$). Ensuite, trouvons un numéro n_1 de façon à avoir

$$p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} < \alpha \leq p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1-1}.$$

Ensuite, trouvons un numéro n_2 de façon à avoir

$$\begin{aligned} p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + p_{n_1+1} + \dots + p_{n_2-1} &\leq \beta < \\ < p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + p_{n_1+1} + \dots + p_{n_2}, \end{aligned}$$

puis un numéro m_2 tel que

$$\begin{aligned} p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + p_{n_1+1} + \dots + p_{n_2} - q_{m_1+1} - \dots - q_{m_2} < \\ < \alpha \leq p_1 + \dots + p_{n_1} - q_1 - \dots - q_{m_1} + p_{n_1+1} + \dots + p_{n_2} - q_{m_1+1} - \dots - q_{m_2-1}. \end{aligned}$$

Continuons la construction indéfiniment. Nous aboutissons à une série

$$b_1 + b_2 + \dots + b_n + \dots \quad (3)$$

qui représente une permutation de la série 6.35 (1). Montrons que la série (3) satisfait à la condition demandée.

a. Les points α et β sont des valeurs d'adhérence de la suite des sommes partielles de la série (3).

En effet, les sommes partielles de la série (3) de numéros $n_1, n_1 + m_1 + n_2, \dots, n_1 + m_1 + n_2 + \dots + m_{h-1} + n_h, \dots$ diffèrent de β au plus de $p_{n_1}, p_{n_2}, \dots, p_{n_h}, \dots$ respectivement, ces dernières quantités tendant vers 0; d'une façon analogue, les sommes partielles de la série (3) de numéros $n_1 + m_1, n_1 + m_1 + n_2 + m_2, \dots, n_1 + m_1 + n_2 + m_2 + \dots + n_h + m_h, \dots$ diffèrent de α au plus de $q_{m_1}, q_{m_2}, \dots, q_{m_h}, \dots$ respectivement, ces dernières quantités tendant elles aussi vers 0.

b. Aucune valeur d'adhérence de la suite des sommes partielles de la série (3) n'appartient à l'extérieur de l'intervalle $[\alpha, \beta]$.

En effet, si, par exemple, $\gamma > \beta$ et $\gamma - \beta = h > 0$, nous trouvons un numéro n_h tel que $p_{n_h} < \frac{h}{2}$; alors toutes les sommes partielles de la série (3) de numéros supérieurs à $n_1 + m_1 + \dots + m_{h-1} + n_h$ ne dépassent pas $\beta + p_{n_h} < \beta + \frac{h}{2}$ et diffèrent de γ d'au moins $\frac{h}{2}$.

Il est évident que les propositions a et b démontrent le théorème en cas de α et β finis. Lorsque l'un des nombres α et β (ou les deux) est infini, la construction doit être un peu modifiée; nous proposons au lecteur de la réaliser.

En particulier, posons $\alpha = \beta = p$, la série (3) converge alors vers le nombre p . Nous voyons qu'en permutant les termes d'une série semi-convergente d'une manière convenable, on peut la faire converger vers une somme donnée à l'avance.

§ 6.4. Séries de vecteurs

6.41. La définition de la convergence d'une série de vecteurs de l'espace R_n est complètement analogue à la définition 6.11 de la convergence d'une série numérique, notamment:

Soit $a_1, \dots, a_m, \dots$ une suite de vecteurs de l'espace R_n . Ecrivons les *sommes partielles*

$$\begin{aligned} s_1 &= a_1 \in R_n, \\ s_2 &= a_1 + a_2 \in R_n, \\ &\dots \dots \dots \\ s_m &= a_1 + a_2 + \dots + a_m \in R_n, \\ &\dots \dots \dots \end{aligned}$$

Si la suite $s_1, \dots, s_m, \dots$ des sommes partielles converge dans l'espace R_n , nous dirons que la série vectorielle

$$a_1 + a_2 + \dots + a_m + \dots \quad (1)$$

converge et a pour somme le vecteur $s = \lim_{n \rightarrow \infty} s_n$. Si, au contraire, la suite de vecteurs $s_1, \dots, s_m, \dots$ diverge, nous dirons que la série (1) diverge.

6.42. Puisque la convergence dans l'espace R_n se réduit à la convergence de chaque coordonnée (3.32.f), on peut dire que la *série des vecteurs*

$$\begin{aligned} a_1 &= (a_{11}, \dots, a_{1n}), \quad a_2 = (a_{21}, \dots, a_{2n}), \dots \\ &\dots, \quad a_m = (a_{m1}, \dots, a_{mn}), \dots \end{aligned}$$

converge si, et seulement si, toutes les séries numériques

$$\begin{aligned} a_{11} + a_{21} + \dots + a_{m1} + \dots, \\ a_{12} + a_{22} + \dots + a_{m2} + \dots, \\ \dots \dots \dots \\ a_{1n} + a_{2n} + \dots + a_{mn} + \dots \end{aligned}$$

convergent; de plus, la somme de la série $\sum a_m$ est le vecteur $p = (p_1, \dots, p_n)$ dont toute composante p_h est la somme de la série numérique correspondante $\sum_{m=1}^{\infty} a_{mh}$.

6.43. Le critère de Cauchy étant valable pour les fonctions vectorielles, en particulier pour les suites vectorielles (4.75), il l'est également pour les séries vectorielles:

Critère de Cauchy pour une série vectorielle. Une série vectorielle $a_1 + a_2 + \dots + a_m + \dots$ ($a_m \in R_n$) converge si, et seulement si, pour tout $\varepsilon > 0$, il existe un nombre N tel que

$$|a_{m+1} + \dots + a_n| < \varepsilon,$$

dès que $n > m \geq N$.

6.44. En particulier, pour toute série vectorielle convergente $\sum_1^\infty a_n$, on a $\lim_{n \rightarrow \infty} a_n = 0$.

6.45. En appliquant le critère de Cauchy de la même façon que dans 6.21, on obtient la condition suffisante de convergence que voici:

Une fois la série numérique

$$|a_1| + |a_2| + \dots + |a_m| + \dots \quad (1)$$

converge, la série vectorielle

$$a_1 + a_2 + \dots + a_m + \dots \quad (2)$$

converge elle aussi.

Une série (2) pour laquelle la série numérique (1) converge est dite *absolument convergente*. Lorsque la série (1) diverge, et la série (2) converge, la dernière est dite *semi-convergente*.

6.46. Les résultats 6.32.a et 6.32.b sur l'addition de séries et la multiplication d'une série par un nombre restent valables pour les séries vectorielles. La proposition 6.32.c, dans le cas général, n'a plus de sens, la multiplication de vecteurs n'étant pas définie. On conserve les résultats 6.33 et 6.34.b sur le groupement de termes d'une série, ainsi que 6.36 (la permutation de termes d'une série). Les énoncés des théorèmes inverses, y compris la réciproque du théorème de Riemann, prennent des formes assez originales; ils sont donnés dans les exercices 17-22.

Les séries à termes complexes considérées comme séries vectorielles dans R_2 rentrent dans le schéma ci-dessus. Pour elles, outre les résultats généraux signalés plus haut, on conserve le théorème sur la multiplication de séries (6.32.c):

Si deux séries à termes complexes

$$a_1 + a_2 + \dots + a_m + \dots, \quad (1)$$

$$b_1 + b_2 + \dots + b_m + \dots \quad (2)$$

convergent et que l'une d'elles au moins converge absolument, alors la série

$$a_1 b_1 + (a_1 b_2 + a_2 b_1) + (a_1 b_3 + a_2 b_2 + a_3 b_1) + \dots \quad (3)$$

converge elle aussi et sa somme vaut le produit des sommes des séries (1) et (2).

La démonstration reproduit celle du théorème 6.32.c.

6.47. Pour des séries vectorielles semi-convergentes, on peut se servir du critère d'Abel-Dirichlet qui est basé sur une transformation spéciale de sommes vectorielles finies appelée *transformation d'Abel*.

a. Soient des nombres $\alpha_1, \dots, \alpha_m$ et des vecteurs $b_1, \dots, b_m$. Posons

$$\alpha'_1 = \alpha_2 - \alpha_1, \alpha'_2 = \alpha_3 - \alpha_2, \dots, \alpha'_{m-1} = \alpha_m - \alpha_{m-1},$$

$$B_1 = b_1, B_2 = b_1 + b_2, \dots, B_m = b_1 + \dots + b_m.$$

Théorème (transformation d'Abel). *Quels que soient les nombres $\alpha_1, \dots, \alpha_m$ et les vecteurs $b_1, \dots, b_m$, l'égalité suivante a lieu :*

$$\sum_{h=1}^m \alpha_h b_h = \alpha_m B_m - \sum_{h=1}^{m-1} \alpha'_h B_h \quad (m=2, 3, \dots). \quad (1)$$

La démonstration se fait par récurrence. Pour $m=2$, l'égalité (1) est, évidemment, remplie :

$$\alpha_1 b_1 + \alpha_2 b_2 = \alpha_2 (b_1 + b_2) - (\alpha_2 - \alpha_1) b_1.$$

Supposons qu'elle soit remplie pour un numéro m ; montrons qu'il en est de même du numéro suivant $m+1$. Par hypothèse de récurrence, nous avons

$$\sum_{h=1}^{m+1} \alpha_h b_h = \sum_{h=1}^m \alpha_h b_h + \alpha_{m+1} b_{m+1} = \alpha_m B_m - \sum_{h=1}^{m-1} \alpha'_h B_h + \alpha_{m+1} b_{m+1}.$$

C'est l'égalité suivante qui est à démontrer :

$$\alpha_m B_m - \sum_{h=1}^{m-1} \alpha'_h B_h + \alpha_{m+1} b_{m+1} = \alpha_{m+1} B_{m+1} - \sum_{h=1}^{m-1} \alpha'_h B_h - \alpha'_m B_m. \quad (2)$$

En réduisant dans (2) les termes $\sum_{h=1}^{m-1} \alpha'_h B_h$, on obtient

$$\begin{aligned} \alpha_m B_m + \alpha_{m+1} b_{m+1} &= \alpha_{m+1} B_{m+1} - \alpha'_m B_m = \\ &= \alpha_{m+1} B_{m+1} - (\alpha_{m+1} - \alpha_m) B_m. \end{aligned}$$

En réduisant d'abord $\alpha_m B_m$ et en divisant ensuite par α_{m+1} (si $\alpha_{m+1} \neq 0$), nous aboutissons à l'égalité $b_{m+1} = B_{m+1} - B_m$ qui est, évidemment, juste. En effectuant les transformations dans l'ordre inverse nous obtenons (2).

b. Signalons une inégalité résultant de (1): si $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_m \geq 0$ et $|B_k| \leq C$ ($k = 1, \dots, n$), alors

$$\left| \sum_{k=1}^m \alpha_k b_k \right| \leq 2C\alpha_1. \quad (3)$$

En effet, dans l'hypothèse formulée, on a $|\alpha_m B_m| \leq \alpha_m C \leq \alpha_1 C$, $\alpha_k \leq 0$ et

$$\begin{aligned} \left| \sum_1^{m-1} \alpha_k B_k \right| &= \left| \sum_1^{m-1} (-\alpha_k) B_k \right| \leq C \sum_1^{m-1} (-\alpha_k) = \\ &= C [(\alpha_1 - \alpha_2) + \dots + (\alpha_{m-1} - \alpha_m)] = C (\alpha_1 - \alpha_m) \leq C\alpha_1; \end{aligned}$$

c'est pourquoi (3) résulte de (1).

c. Critère d'Abel-Dirichlet pour les séries vectorielles. La série

$$\sum_1^{\infty} \alpha_m b_m \quad (4)$$

converge si $\alpha_m \searrow 0$ et les vecteurs $B_m = b_1 + \dots + b_m$ sont bornés, $|B_m| \leq C$.

Démonstration. Appliquons l'estimation (3) à un morceau de la série (4). Nous avons

$$\left| \sum_{k=p}^q \alpha_k b_k \right| \leq 2C_1 \alpha_p,$$

ù

$$\begin{aligned} C_1 &= \sup_{p \leq r \leq q} |b_p + \dots + b_r| = \sup_{p \leq r \leq q} |(b_1 + \dots + b_r) - \\ &\quad - (b_1 + \dots + b_{p-1})| = \sup_{p \leq r \leq q} |B_r - B_{p-1}| \leq 2C. \end{aligned}$$

La quantité obtenue tend vers zéro pour $p \rightarrow \infty$ et n'importe quel $q > p$. On peut donc appliquer le critère de Cauchy 6.43 à la série (4), de sorte que la série converge.

d. Le critère de Leibniz 6.23 est un cas particulier du critère d'Abel-Dirichlet. On le verra en posant $n = 1$ et $B_1 = 1$, $B_2 = -1$, $B_3 = 1$, $B_4 = -1$, ... Avec la seule condition $n = 1$, le critère d'Abel-Dirichlet est déjà plus puissant que celui de Leibniz.

Considérons deux séries réelles ($n = 1$)

$$\sum_0^{\infty} \alpha_m \sin m\theta, \quad (5)$$

$$\sum_0^{\infty} \alpha_m \cos m\theta \quad (6)$$

et la série de nombres complexes

$$\sum_{m=0}^{\infty} \alpha_m (\cos m\theta + i \sin m\theta). \quad (7)$$

Trouvons les valeurs de θ pour lesquelles les séries données convergent. Posons $z = \cos \theta + i \sin \theta$; alors (5.72) $|z| = 1$ et

$$z^m = \cos m\theta + i \sin m\theta.$$

Ensuite

$$\begin{aligned} \sum_{k=0}^m (\cos k\theta + i \sin k\theta) &= \sum_{k=0}^m z^k = \frac{1 - z^{m+1}}{1 - z}, \\ \left| \sum_{k=0}^m (\cos k\theta + i \sin k\theta) \right| &\leq \frac{2}{|1 - z|} = \frac{2}{|1 - \cos \theta - i \sin \theta|} = \\ &= \frac{2}{\sqrt{(1 - \cos \theta)^2 + \sin^2 \theta}} = \sqrt{\frac{2}{1 - \cos \theta}}. \end{aligned} \quad (8)$$

On voit que la quantité

$$\left| \sum_{k=0}^m (\cos k\theta + i \sin k\theta) \right|$$

reste bornée pour $m \rightarrow \infty$ si $\theta \neq 0, \pm 2\pi, \pm 4\pi, \dots$. Supposons maintenant que les nombres α_m tendent vers zéro en décroissant. On peut alors appliquer à la série (7) le critère d'Abel-Dirichlet; on en conclut que la série (7) converge lorsque $\alpha_m \searrow 0$ et $\theta \neq 0, \pm 2\pi, \dots$. Dans les mêmes conditions, les séries (5) et (6) convergent elles aussi puisqu'elles représentent les parties réelle et imaginaire de la série (7). La série (5) converge même pour $\theta = 0, \pm 2\pi, \dots$, car ces valeurs de θ annulent tous ses termes.

e. L'inégalité (8) implique les inégalités analogues dans le domaine réel:

$$\begin{aligned} \left| \sum_{k=0}^m \cos k\theta \right| &\leq \sqrt{\frac{2}{1 - \cos \theta}}, \\ \left| \sum_{k=0}^m \sin k\theta \right| &\leq \sqrt{\frac{2}{1 - \cos \theta}}. \end{aligned} \quad (9)$$

Ces estimations sont souvent employées lorsqu'on a affaire au critère d'Abel-Dirichlet.

f. Dans le cas particulier où $\alpha_m = r^m$ ($r < 1$), il est aisé de trouver les sommes des séries (5)-(7). Notamment, en sommant la

progression géométrique on a

$$\sum_0^{\infty} z^m = \sum_0^{\infty} r^m (\cos \theta + i \sin \theta)^m = \frac{1}{1 - r(\cos \theta + i \sin \theta)}.$$

En séparant les parties réelle et imaginaire, on trouve

$$\begin{aligned} \frac{1}{1 - r(\cos \theta + i \sin \theta)} &= \frac{1 - r(\cos \theta - i \sin \theta)}{[1 - r(\cos \theta + i \sin \theta)][1 - r(\cos \theta - i \sin \theta)]} = \\ &= \frac{1 - r \cos \theta + i r \sin \theta}{1 - 2r \cos \theta + r^2}, \end{aligned}$$

d'où

$$\sum_0^{\infty} r^m \cos m\theta = \operatorname{Re} \sum_0^{\infty} r^m e^{im\theta} = \frac{1 - r \cos \theta}{1 - 2r \cos \theta + r^2},$$

$$\sum_0^{\infty} r^m \sin m\theta = \operatorname{Im} \sum_0^{\infty} r^m e^{im\theta} = \frac{r \sin \theta}{1 - 2r \cos \theta + r^2} \quad (\text{voir 8.63(5)}).$$

6.48. Séries bilatérales.

a. Soit une suite « infinie dans les deux sens » de vecteurs $\dots, a_{-n}, \dots, a_{-1}, a_0, a_1, \dots, a_n, \dots$. Nous définissons la somme de la « série bilatérale »

$$\sum_{-\infty}^{\infty} a_h \quad (1)$$

comme limite des sommes partielles

$$\sum_{-m}^n a_h,$$

lorsque m, n tendent vers l'infini indépendamment l'un de l'autre, à condition que ladite limite existe. Autrement dit, la série (1) converge et a s pour somme si, pour tout $\varepsilon > 0$, il existe un numéro N tel que, quels que soient $m, n > N$, on a l'inégalité

$$\left| s - \sum_{-m}^n a_h \right| < \varepsilon. \quad (2)$$

b. Le théorème suivant ramène la convergence d'une série bilatérale à celle de séries unilatérales.

Théorème. La série (1) converge si, et seulement si, les deux séries « unilatérales »

$$\sum_0^{\infty} a_h \quad \text{et} \quad \sum_1^{\infty} a_{-h} \quad (3)$$

convergent ; si les deux dernières séries convergent et

$$\sum_0^{\infty} a_k = A, \quad \sum_{-1}^{\infty} a_{-k} = B,$$

alors la somme s de la série (1) vaut $A + B$.

Démonstration. Supposons que les séries (3) convergent. Pour tout $\varepsilon > 0$, on peut trouver un numéro N tel que si $n > N$, $m > N$, alors

$$\left| A - \sum_0^n a_k \right| < \frac{\varepsilon}{2}, \quad \left| B - \sum_1^m a_{-k} \right| < \frac{\varepsilon}{2},$$

d'où

$$\left| (A + B) - \sum_{-m}^n a_k \right| = \left| (A - \sum_0^n a_k) + (B - \sum_1^m a_{-k}) \right| < \varepsilon,$$

ce qui veut dire que la série (1) converge et a $A + B$ pour somme.

Réciproquement, si c'est la série (1) qui converge, alors il résulte de l'inégalité (2) que, pour n'importe quels $p > N$, $q > 0$,

$$\left| \sum_p^{p+q} a_k \right| = \left| (s - \sum_{-m}^{p-1} a_k) - (s - \sum_{-m}^{p+q} a_k) \right| < 2\varepsilon,$$

ce qui implique, en vertu du critère de Cauchy 6.43, la convergence de la série $\sum_0^{\infty} a_k$. La convergence de la série $\sum_1^{\infty} a_{-k}$ se démontre de façon analogue, ce qui achève la démonstration du théorème.

c. Exemple. La série bilatérale

$$\sum_{-\infty}^{\infty} c_n (\cos nt + i \sin nt) = \sum_{-\infty}^{\infty} c_n e^{int} \text{ (voir 8.63(5))}$$

converge partout pour $t \in R$, si $\sum_{-\infty}^{\infty} |c_n| < \infty$; elle converge partout, sauf, peut-être aux points $t = 0, \pm 2\pi, \pm 4\pi, \dots$, si $c_n \searrow 0$, $c_{-n} \searrow 0$ ($n \rightarrow +\infty$) (cf. 6.47.d).

6.49.a. Sommation symétrique de séries bilatérales. Une série vectorielle bilatérale

$$\sum_{-\infty}^{\infty} a_k \tag{1}$$

est dite *symétriquement sommable* s'il existe la limite pour $n \rightarrow \infty$ de ses sommes partielles symétriques

$$s_n = \sum_{-n}^n a_k, \quad n = 0, 1, 2, \dots; \quad (2)$$

dans ce cas, la limite s des sommes (2) s'appelle *somme symétrique* de la série (1).

Si la série (1) converge au sens de 6.48, alors, évidemment, elle est symétriquement sommable et sa somme symétrique se confond avec sa somme ordinaire. Mais une série symétriquement sommable ne converge pas forcément au sens de 6.48. Par exemple, la série bilatérale composée des nombres a_k valant 1 pour $k > 0$, 0 pour $k = 0$ et -1 pour $k < 0$, est bien symétriquement sommable (et sa somme symétrique vaut 0) sans être convergente au sens de 6.48.

b. La sommabilité symétrique de séries bilatérales se ramène elle aussi à la convergence ordinaire :

T h é o r è m e. *La série bilatérale (1) est symétriquement sommable si, et seulement si, la série unilatérale*

$$a_0 + \sum_{k=1}^{\infty} (a_k + a_{-k}) \quad (3)$$

converge; la série (3) étant convergente, sa somme se confond avec la somme symétrique de la série (1).

D é m o n s t r a t i o n. Le théorème découle de la coïncidence évidente de la n -ième somme partielle de la série (3) avec la n -ième somme partielle symétrique (2) de la série (1).

La sommation symétrique est utilisée, par exemple, dans la théorie des séries trigonométriques de la forme

$$\sum_{-\infty}^{\infty} c_n e^{int}$$

(III^e partie, chapitre 14).

c. **E x e m p l e.** La série bilatérale

$$\dots - \frac{e^{-int}}{n} - \dots - \frac{e^{-2it}}{2} - \frac{e^{-it}}{1} + \frac{e^{it}}{1} + \frac{e^{2it}}{2} + \dots + \frac{e^{nit}}{n} + \dots$$

converge pour tous les $t \neq 0, \pm 2\pi, \pm 4\pi, \dots$ (6.48.c), mais ne converge pas pour $t = 0, \pm 2\pi, \pm 4\pi, \dots$, car, pour ces dernières valeurs de t , l'hypothèse du théorème 6.48.h n'est pas remplie.

Tout de même, elle est symétriquement sommable pour toutes les valeurs de $t \in R_1$, parce que la série

$$\sum_{k=1}^{\infty} \frac{e^{ikt} - e^{-ikt}}{k} = 2t \sum_{k=1}^{\infty} \frac{\sin kt}{k}$$

est partout convergente (6.47.d).

§ 6.5. Séries de fonctions

6.51. Nous avons déjà rencontré des séries de fonctions (§ 6.4, par ex.). Donnons à présent quelques définitions et théorèmes généraux se rapportant à la convergence de telles séries. Soit une série

$$a_1(x) + a_2(x) + \dots + a_n(x) + \dots \quad (1)$$

de fonctions définies sur un ensemble quelconque E et à valeurs dans R_m . Conformément à 5.91 et 5.93, formulons les définitions de la convergence et de la convergence uniforme d'une série (1).

a. Définition. Si la suite des sommes partielles de la série (1)

$$\begin{aligned} s_1(x) &= a_1(x), \\ s_2(x) &= a_1(x) + a_2(x), \\ &\dots \dots \dots \\ s_n(x) &= a_1(x) + \dots + a_n(x), \\ &\dots \dots \dots \end{aligned}$$

converge sur l'ensemble E , nous disons que la série (1) converge sur l'ensemble E et appelons somme de la série (1) la limite des fonctions $s_n(x)$.

b. Définition. Si la suite $s_n(x)$ converge uniformément sur l'ensemble E , nous disons que la série (1) converge uniformément sur l'ensemble E .

6.52. Le critère de Cauchy 5.98 est formulé de la façon suivante pour une série de fonctions vectorielles:

Pour que la série 6.51 (1) soit uniformément convergente sur l'ensemble E , il faut et il suffit que, pour tout $\varepsilon > 0$, il existe un numéro N tel que, quels que soient $n \geq N$, $p > n$, $x \in E$, on ait

$$|a_{n+1}(x) + \dots + a_p(x)| \leq \varepsilon.$$

6.53. Donnons encore une condition suffisante simple de convergence uniforme de la série 6.51 (1):

Critère de Weierstrass. Si

$$\sup_{x \in E} |a_n(x)| = \alpha_n$$

et la série $\sum_1^{\infty} \alpha_n$ converge, alors la série (1) converge uniformément.

Cette proposition résulte de 6.52, vu l'inégalité

$$|a_{n+1}(x) + \dots + a_p(x)| \leq \alpha_{n+1} + \dots + \alpha_p$$

et le critère de Cauchy pour une série numérique (6.11.c).

6.54. Les théorèmes suivants sont des conséquences immédiates des théorèmes 5.94-5.95.

a. **Théorème.** Si la série 6.51 (1) est formée de fonctions bornées à valeurs dans R_m et si elle converge uniformément sur E , alors sa somme $s(x)$ est bornée sur E .

b. **Théorème.** Si la série 6.51 (1) est formée de fonctions continues sur un espace métrique E , à valeurs dans R_m , et si elle converge uniformément sur E , alors la somme de la série (1) est aussi une fonction continue sur l'espace E , à valeurs dans R_m .

§ 6.6. Séries de puissances

6.61. Les séries de puissances sont d'un intérêt particulier dans le domaine complexe.

Définition. Une série de la forme

$$a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \dots + a_n(z - z_0)^n + \dots, \quad (1)$$

où $z = x + iy$, $z_0 = x_0 + iy_0$, et les coefficients $a_0, a_1, \dots$ sont des nombres complexes, s'appelle *série de puissances* (ou *série entière*).

6.62. Il est utile de déterminer le domaine de convergence de cette série, c'est-à-dire l'ensemble des valeurs de z pour lesquelles la série 6.61 (1) converge. On a eu l'occurrence le théorème suivant :

Théorème (Cauchy-Hadamard). Posons

$$\frac{1}{R} = \overline{\lim} \sqrt[n]{|a_n|}$$

où la limite supérieure est considérée sur la droite numérique achevée. La série 6.61 (1) converge, et même absolument, en tout point z avec $|z - z_0| < R$ et diverge en tout point z avec $|z - z_0| > R$.

Remarque. Si $\overline{\lim} \sqrt[n]{|a_n|} = \infty$, on pose $R = 0$. Si $\overline{\lim} \sqrt[n]{|a_n|} = 0$, on pose $R = \infty$, et alors le théorème affirme que la série 6.61 (1) converge pour tout $z \in C$.

Démonstration. Abordons le cas d'un $R \neq 0$ fini. Soit

$$|z - z_0| < \frac{1}{\lim \sqrt[n]{|a_n|}}.$$

Cette inégalité peut être mise sous la forme

$$|z - z_0| \lim \sqrt[n]{|a_n|} = \lim \sqrt[n]{|a_n|} |z - z_0|^n < 1,$$

d'où, en vertu du critère de Cauchy 6.14.b et 6.45, la convergence de la série 6.61 (1).

Dans le cas où

$$|z - z_0| > \frac{1}{\lim \sqrt[n]{|a_n|}},$$

on a de même

$$|z - z_0| \lim \sqrt[n]{|a_n|} = \lim \sqrt[n]{|a_n (z - z_0)^n|} > 1,$$

et, d'après le même critère 6.14.b et 6.44, la série 6.61 (1) diverge.

Laissons à nos lecteurs le soin d'examiner les cas $R = 0$ et $R = \infty$.

6.63. Le théorème 6.62 montre que le domaine $G \in R_1$ de convergence de la série de puissances 6.61 (1) est le cercle de rayon R , et c'est la raison pour laquelle le nombre R s'appelle *rayon de convergence de la série* 6.61 (1). Plus précisément, le domaine G contient tous les points intérieurs de ce cercle (les z pour lesquels $|z - z_0| < R$) et ne contient aucun point extérieur (pour lequel $|z - z_0| > R$). En ce qui concerne les points de la circonférence même $|z - z_0| = R$, différents cas peuvent se présenter.

Exemples. a. La série $\sum_1^{\infty} \frac{z^n}{n^n}$ converge pour tous les z , car

$$\frac{1}{R} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n^n}} = \lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

b. La série $\sum_1^{\infty} n^n z^n$ ne converge que pour $z = 0$, puisque

$$\frac{1}{R} = \lim \sqrt[n]{n^n} = \lim n = \infty.$$

c. Pour tout α fixe, la série $\sum_1^{\infty} n^{\alpha} z^n$ a 1 pour rayon de convergence, vu que, en vertu de 5.58 (6),

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n^{\alpha}}} = \lim_{n \rightarrow \infty} \left(\frac{1}{n^{\frac{\alpha}{n}}} \right) = 1.$$

Mais, pour $\alpha \geq 0$, la série ne converge en aucun point de la circonférence $|z| = 1$, ses termes ne tendant pas vers 0. Pour $\alpha < -1$, la série converge, et même absolument, en tout point de la circonférence $|z| = 1$ (6.15.a). Si $-1 \leq \alpha < 0$, alors la série diverge pour $z = 1$ tout en convergeant pour tous les $z \neq 1$, $|z| = 1$ (cf. 6.47.d).

6.64. Soit une série de puissances 6.61 (1) ayant R pour rayon de convergence. Posons la question suivante: est-ce que cette série converge uniformément dans le cercle $|z - z_0| < R$? Dans le cas général, la réponse est négative; ainsi, si la série

$$1 + z + z^2 + \dots$$

convergeait uniformément dans son cercle de convergence $|z| < 1$, alors sa somme serait bornée (6.54.a); or, sa somme vaut $\frac{1}{1-z}$ (6.11.b) et n'est pas bornée dans le cercle $|z| < 1$.

Tout de même, le théorème suivant a lieu:

a. La série 6.61 (1) converge uniformément dans tout cercle $|z - z_0| \leq R_1 < R$.

En effet, pour $|z - z_0| \leq R_1$, on a l'inégalité suivante

$$|a_n (z - z_0)^n| \leq |a_n| R_1^n,$$

et la série des nombres $|a_n| R_1^n$ converge (on l'a vu en démontrant le théorème de Cauchy-Hadamard 6.62). Donc, on peut appliquer le critère de Weierstrass 6.53 au cercle $|z| \leq R_1$, ce qui conduit au résultat cherché.

b. Une conséquence de a et de 6.54.a, b:

La somme $s(z)$ de la série 6.61 (1) est une fonction bornée et continue dans tout cercle $|z - z_0| \leq R_1$.

La fonction $s(z)$ est donc continue dans tout le cercle $|z - z_0| < R$, chaque point de ce cercle appartenant à un cercle plus petit $|z - z_0| \leq R_1$. Mais elle n'est pas forcément bornée dans tout le cercle $|z - z_0| < R$, ce qu'on a vu plus haut.

6.65. Nous savons qu'une série de puissances peut converger aussi bien que diverger sur la frontière de son cercle de convergence. Mais si elle converge en un point z_1 de la frontière, tout l'intervalle fermé qui joint le centre du cercle au point z_1 peut être inclus dans le domaine de convergence uniforme de la série. Pour le démontrer, il suffit de considérer le cas où $z_0 = 0$, $z_1 = t_1 > 0$.

Théorème d'Abel. Si une série de puissances $\sum_0^{\infty} a_n t^n$ converge en un point $t_1 > 0$, alors elle converge uniformément sur l'intervalle $0 \leq t \leq t_1$.

Démonstration. Pour un $\varepsilon > 0$ donné, nous avons à trouver un N tel que l'inégalité $\left| \sum_{p+1}^q a_n t^n \right| < \varepsilon$ ait lieu pour tous les $|t| \leq t_1$, dès que $p \geq N$, $q \geq N$. Appliquons la transformation d'Abel 6.47 (1) en posant $\alpha_k = \left(\frac{t}{t_1}\right)^k$, $b_k = a_k t_1^k$ ($k = p+1, \dots, q$, $p < q$ étant pour le moment arbitraires) :

$$\begin{aligned} \sum_{p+1}^q a_n t^n &= \sum_{p+1}^q a_n t_1^n \left(\frac{t}{t_1}\right)^n = \left(\frac{t}{t_1}\right)^q \sum_{p+1}^q a_n t_1^n - \\ &- \sum_{p+1}^{q-1} \left[\left(\frac{t}{t_1}\right)^{n+1} - \left(\frac{t}{t_1}\right)^n \right] \sum_{h=p+1}^n a_h t_1^h. \end{aligned}$$

Choisissons N de manière à avoir $\left| \sum_{p+1}^n a_h t_1^h \right| < \frac{\varepsilon}{2}$ dès que $p \geq N$, $n \geq N$; nous avons alors

$$\begin{aligned} \left| \sum_{p+1}^q a_n t^n \right| &\leq \frac{\varepsilon}{2} \left(\frac{t}{t_1}\right)^q + \frac{\varepsilon}{2} \max \sum_{p+1}^{q-1} \left| \left(\frac{t}{t_1}\right)^{n+1} - \left(\frac{t}{t_1}\right)^n \right| \leq \\ &\leq \frac{\varepsilon}{2} \cdot 1 + \frac{\varepsilon}{2} \max \left[\left(\frac{t}{t_1}\right)^{p+1} - \left(\frac{t}{t_1}\right)^{q+1} \right] \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} - \varepsilon, \end{aligned}$$

ce qu'il fallait démontrer.

6.66. Exemples.

a. La série

$$1 + \frac{z}{1^\alpha} + \frac{z^2}{2^\alpha} + \dots + \frac{z^n}{n^\alpha} + \dots \quad (0 < \alpha \leq 1)$$

converge pour tous les $|z| = 1$, $z \neq 1$ (6.47.d). D'après le théorème d'Abel 6.65, elle converge uniformément sur tout rayon du cercle $|z| \leq 1$ aboutissant à un point z_0 tel que $|z_0| = 1$, $z_0 \neq 1$ (ceci ne vaut nullement dire que la série converge uniformément sur l'ensemble des points de tous ces rayons !).

b. La série

$$\sum_{n=1}^{\infty} \frac{\alpha(\alpha+1) \dots (\alpha+n-1)}{\beta(\beta+1) \dots (\beta+n-1)} z^{n-1}$$

(α, β différents de 0, -1, -2, ...) converge pour $z = 1$ si $\beta > \alpha + 1$ (6.17.b), donc pour $|z| < 1$, et converge uniformément sur l'intervalle $0 \leq x \leq 1$. Si $\beta > \alpha$, la même série converge pour

$z = -1$ (6.24.c), donc converge uniformément sur l'intervalle $-1 \leq x \leq 0$.

6.67. Retenons la conséquence suivante du théorème d'Abel :

Théorème. *Supposons qu'une fonction $f(x)$ soit continue pour $0 \leq x \leq x_1$ et que l'égalité*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

ait lieu pour $0 \leq x < x_1$. Alors, si la série

$$\sum_{n=0}^{\infty} a_n x_1^n \quad (1)$$

converge, on a

$$f(x_1) = \sum_{n=0}^{\infty} a_n x_1^n.$$

Démonstration. Pour tous les $x \in [0, x_1]$, posons

$$\sum_{n=0}^{\infty} a_n x^n = s(x). \quad (2)$$

La série (1) étant convergente, la série (2) converge uniformément sur $[0, x_1]$ d'après le théorème d'Abel. Donc, en vertu de 6.65.b, la fonction $s(x)$ est continue sur $[0, x_1]$. La fonction $f(x)$ étant elle aussi continue sur $[0, x_1]$ par hypothèse, et vu que $f(x) = s(x)$ pour $0 \leq x < x_1$, on a

$$f(x_1) = \lim_{x \rightarrow x_1} f(x) = \lim_{x \rightarrow x_1} s(x) = s(x_1) = \sum_{n=0}^{\infty} a_n x_1^n,$$

ce qu'il fallait démontrer.

Exercices

1. Soient $a_n > 0$ ($n=1, 2, \dots$) et $\frac{a_{n+1}}{a_n} = 1 - \frac{1}{n} + \frac{\beta_n}{n^2}$, où les nombres $|\beta_n|$ sont bornés et $|\beta_n| \leq C$. Démontrer que la série $\sum_1^{\infty} a_n$ diverge (Gauss).

2. La série hypergéométrique de Gauss (α, β, γ sont différents des nombres $0, -1, -2, \dots$)

$$F(\alpha, \beta, \gamma, x) = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha+1) \dots (\alpha+n-1) \beta(\beta+1) \dots (\beta+n-1)}{n! \gamma(\gamma+1) \dots (\gamma+n-1)} x^n$$

converge absolument pour $|x| < 1$, diverge pour $|x| > 1$; si $x = 1$, elle converge (absolument) pour $\gamma > \alpha + \beta$, diverge pour $\gamma \leq \alpha + \beta$; si $x = -1$, elle est absolument convergente pour $\gamma > \alpha + \beta$, semi-convergente pour $-1 < \gamma - (\alpha + \beta) \leq 0$, divergente pour $\gamma - (\alpha + \beta) \leq -1$.

3. Si $a_n \geq a_{n+1} > 0$ ($n=1, 2, \dots$) et la série $\sum_1^\infty a_n$ converge, alors a_n est un infiniment petit par rapport à $\frac{1}{n}$.

Remarque. Il n'existe pas de fonction $\varphi(n)$ tendant vers 0 plus vite que $\frac{1}{n}$ que l'on pourrait substituer à $\frac{1}{n}$ dans l'énoncé (A. Némirovski).

4. Soit $\sum_1^\infty a_n$ une série convergente ($a_n > 0$); démontrer qu'il existe une suite $b_1 \leq b_2 \leq \dots \leq b_n \leq \dots$, $\lim b_n = \infty$, telle que la série $\sum_1^\infty a_n b_n$ converge elle aussi.

5. Montrer que le produit de la série convergente

$$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \dots$$

par elle-même est une série divergente.

6. Démontrer que pour une série de Leibniz 6.23, la différence entre la somme de la série et la somme de ses n premiers termes ne dépasse pas en valeur absolue le module du $(n+1)$ ième terme.

7. Démontrer que le rayon de convergence de la série de puissances $\sum_0^\infty a_n (x - x_1)^n$, pour laquelle $\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = r$ existe, est $\frac{1}{r}$.

8. Si $a_n \geq 0$ pour tous les $n=1, 2, \dots$ et $f(t) = \sum_0^\infty a_n t^n \leq C$ pour $0 \leq t < 1$, alors, lorsque $t \rightarrow 1$, la limite de la fonction $f(t)$ existe et est égale à $\sum_0^\infty a_n$.

9. Supposons que les nombres $a_{nh} \geq 0$ ($n=1, 2, 3, \dots$, $h=1, 2, \dots$) et b_n ($n=1, 2, \dots$) vérifient les conditions:

1) $a_{nh} \leq b_n$ quels que soient $h=1, 2, \dots$ et $n=1, 2, \dots$,

2) $\sum_{n=1}^\infty b_n < \infty$,

3) $\lim_{h \rightarrow \infty} a_{nh} = a_n$ ($n=1, 2, \dots$).

Posons $\sum_{n=1}^\infty a_{nh} = s_h$ ($h=1, 2, \dots$) et $\sum_{h=1}^\infty a_h = s$. Démontrer que $s = \lim_{h \rightarrow \infty} s_h$.

10. Produits infinis. Soit $z_1, z_2, \dots, z_n, \dots$ une suite de nombres complexes. On dit que le produit infini

$$\prod_{h=1}^\infty z_h = z_1 z_2 \dots z_h \dots$$

converge si la suite numérique $P_n = \prod_{h=1}^n x_h$ ($n = 1, 2, \dots$) converge vers une limite finie $P \neq 0$; le nombre P est appelé valeur du produit $\prod_1^\infty x_h$.

Démontrer que la convergence du produit infini implique la relation $\lim_{n \rightarrow \infty} x_n = 1$.

11. (Suite.) Soit $x_1 > 0, x_2 > 0, \dots, x_n > 0, \dots$; pour que le produit $\prod_{h=1}^\infty x_h$ converge, il faut et il suffit que converge la série $\sum_{h=1}^\infty \log_b x_h$ (de base $b > 1$ quelconque).

12. (Suite.) Posons $x_h = 1 + \omega_h$. Lorsque tous les ω_h sont de même signe, le produit $\prod_{h=1}^\infty (1 + \omega_h)$ converge si, et seulement si, la série $\sum_{h=1}^\infty \omega_h$ converge. (Si les ω_h sont de différents signes, le résultat n'est plus valable; cf. l'exercice 15 pour le chapitre 8.)

13. Démontrer que l'égalité

$$(1+x)(1+x^2)(1+x^4)(1+x^8) \dots (1+x^{2^{n-1}}) \dots = \frac{1}{1-x}$$

a lieu pour $|x| < 1$.

14. Démontrer l'identité d'Euler

$$\prod_{h=1}^\infty \frac{1}{1 - \frac{1}{p_h^x}} = \sum_{n=1}^\infty \frac{1}{n^x}$$

($x > 1$, $p_1, p_2, \dots, p_n, \dots$ est la suite de tous les nombres premiers).

15. Démontrer que la série $\sum_{h=1}^\infty \frac{1}{p_h}$ (p_h sont des nombres premiers) diverge.

16. Soit $\{q_h\}$ la suite de tous les nombres naturels dont la représentation décimale (1.77) ne contient pas le chiffre neuf. La série $\sum_1^\infty \frac{1}{q_h}$, est-elle convergente ou non?

17. On considère une série

$$u_1 + u_2 + \dots + u_m + \dots \quad (1)$$

de vecteurs de l'espace euclidien R_n . Un vecteur $p \in R_n$ est appelé vecteur de convergence absolue de la série (1) si la série $\sum_1^\infty (p, u_m)$ converge absolument.

Démontrer que la série (1) converge absolument si, et seulement si, tout vecteur $p \in R_n$ est un vecteur de convergence absolue de cette série.

18. (Suite.) Un vecteur $q \in R_n$ est appelé vecteur de divergence absolue de la série (1) lorsque les vecteurs de la série (1) qui appartiennent à n'importe quel angle solide contenant le vecteur q forment une série absolument divergente. Démontrer que la série (1) diverge absolument si, et seulement si, elle possède au moins un vecteur de divergence absolue.

19. (Suite.) Soit $k_1 < k_2 < \dots$ une suite de nombres naturels, et soit $f_1 < f_2 < \dots$ la suite des nombres naturels qui ne figurent pas dans la première suite. Les séries $u_{k_1} + u_{k_2} + \dots$ et $u_{f_1} + u_{f_2} + \dots$ sont appelées parties complémentaires de la série (1). Démontrer que si l'une des parties complémentaires d'une série convergente converge, alors il en est de même de l'autre partie, et que toute permutation de termes de la série (1) qui conserve l'ordre de succession de termes dans chaque partie conserve la somme de la série.

20. (Suite.) Si la série (1) est semi-convergente et q est un vecteur de divergence absolue de la série (1), alors il existe une partie $u_{k_1} + \dots$ de la série (1) dont les projections des termes sur le vecteur q forment une série absolument divergente et les composantes orthogonales une série absolument convergente.

21. (Suite.) Si tout vecteur $e \neq 0$ est un vecteur de divergence absolue de la série (1), alors n'importe quel vecteur $f \in R_n$ est la somme d'une série obtenue en permutant les termes de la série (1).

22. (Suite, théorème de Steinitz.) Pour toute série semi-convergente (1), l'ensemble des sommes de toutes les séries obtenues de (1) en permutant ses termes est une variété linéaire dans R_n qui est orthogonale au sous-espace A de tous les vecteurs de convergence absolue et qui passe par la somme des projections des termes de la série (1) sur A (cette somme est définie d'une façon univoque).

Historique

La sommation de progressions géométriques infinies de raison inférieure à 1 s'effectue déjà dans l'Antiquité (Archimède). La divergence de la série harmonique est établie par le savant italien Mengoli en 1650. Les séries de puissances apparaissent chez Newton (1665) qui croit possible de représenter n'importe quelle fonction par une série de puissances. Les séries figurent constamment dans les calculs de savants du XVIII^e siècle, mais ce n'est pas toujours, et de loin, que l'on prête attention au problème de convergence *). Une théorie rigoureuse des séries date des travaux de Gauss (1812), de Bolzano (1817) et, enfin, de Cauchy (1821) qui donne, pour la première fois, la définition actuellement en vigueur de la somme d'une série convergente et établit des théorèmes de base.

*) Ce qui prête plus d'une fois à des spéculations ontoscientifiques. Ainsi, en 1703, le camaldule et théologien italien Grandi tâche de « démontrer » l'existence de Dieu en jonglant avec la série $1 - 1 + 1 - \dots$.

De tous les progrès théoriques, aucun ne passe sans doute pour un triomphe aussi élevé de l'esprit humain que l'invention du calcul infinitésimal.

F. Engels

**DEUXIÈME
PARTIE**

**Calcul différentiel
et intégral**

Dérivée

Pour trouver le maximum ou le minimum d'une quantité $f(x)$, il faut former l'expression $f(x+h) - f(x)$, où h est un nombre indéterminé. Puis, après en avoir chassé les fractions et les radicaux et réduit les termes semblables, il faut diviser l'expression obtenue par ce h indéterminé. En posant, ensuite, $h = 0$ dans les termes qui restent nous avons une équation contenant la lettre x dont les racines sont bien les maxima et les minima.

P. Fermat (1629)

Nous sommes maintenant en mesure de passer à l'une des notions fondamentales de l'analyse mathématique, celle de dérivée.

§ 7.1. Définition d'une dérivée

7.11. Soit une fonction réelle (finie) $y = f(x)$ définie dans un intervalle (a, b) . Soit $x_0 \in (a, b)$; écrivons le rapport

$$\frac{f(x_0+h) - f(x_0)}{h} \quad (x_0 + h \in (a, b)). \quad (1)$$

Lorsque ce rapport possède une limite (finie) pour $h \rightarrow 0$, nous disons que la fonction $f(x)$ est *dérivable* pour $x = x_0$, et nous posons

$$\lim_{h \rightarrow 0} \frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) = [f(x)]'_{x=x_0}. \quad (2)$$

Le nombre $f'(x_0)$ s'appelle *dérivée* de la fonction $f(x)$ pour $x = x_0$.

Géométriquement, le rapport (1) représente le coefficient angulaire de la droite qui coupe le graphique de la fonction $y = f(x)$ aux points d'abscisses x_0 et $x_0 + h$ (fig. 7.1).

L'expression (2) est le coefficient angulaire d'une droite passant par le point (x_0, y_0) ; nous l'appellerons *tangente au graphique de la fonction $f(x)$ au point d'abscisse x_0* .

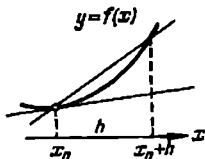


Fig. 7.1.

7.12. Si la fonction $y = f(x)$ possède une dérivée pour $x = x_0$, alors le rapport (1) est borné, lorsque $h \rightarrow 0$, par une constante c ; d'où, quels que soient $|h|$ assez petits,

$$|f(x_0 + h) - f(x_0)| \leq c |h|;$$

par conséquent, la fonction $f(x)$ est continue pour $x = x_0$.

7.13. Règles fondamentales de calcul d'une dérivée. Soient $f(x)$ et $g(x)$ deux fonctions définies dans l'intervalle (a, b) et dérivables pour $x = x_0$. Alors les fonctions $f(x) + g(x)$, $\alpha f(x)$ (α étant n'importe quel nombre réel), $f(x) \cdot g(x)$, $\frac{f(x)}{g(x)}$ sont dérivables pour $x = x_0$ (la dernière dans le cas où $g(x_0) \neq 0$), les égalités suivantes étant valables:

$$(f(x) + g(x))'_{x=x_0} = f'(x_0) + g'(x_0); \quad (1)$$

$$(\alpha f(x))'_{x=x_0} = \alpha f'(x_0); \quad (2)$$

$$(f(x) \cdot g(x))'_{x=x_0} = f'(x_0) g(x_0) + f(x_0) g'(x_0); \quad (3)$$

$$\left(\frac{f(x)}{g(x)} \right)'_{x=x_0} = \frac{g(x_0) f'(x_0) - f(x_0) g'(x_0)}{g^2(x_0)}. \quad (4)$$

Démonstration. (1). Nous avons

$$\frac{f(x_0 + h) + g(x_0 + h) - (f(x_0) + g(x_0))}{h} = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{g(x_0 + h) - g(x_0)}{h},$$

et le résultat découle de 4.36.a.

(2). Nous avons

$$\frac{\alpha f(x_0 + h) - \alpha f(x_0)}{h} = \alpha \frac{f(x_0 + h) - f(x_0)}{h},$$

et le résultat découle de 4.36.b.

(3). Nous avons

$$\begin{aligned} \frac{f(x_0 + h) g(x_0 + h) - f(x_0) g(x_0)}{h} &= \frac{f(x_0 + h) g(x_0 + h) - f(x_0) g(x_0 + h)}{h} + \\ &+ \frac{f(x_0) g(x_0 + h) - f(x_0) g(x_0)}{h} = g(x_0 + h) \frac{f(x_0 + h) - f(x_0)}{h} + \\ &+ f(x_0) \frac{g(x_0 + h) - g(x_0)}{h}, \end{aligned}$$

et le résultat découle de 4.36.b et 4.36.a, compte tenu de 7.12.

Etant donné (3), la démonstration de (4) se réduit au cas où $f(x) \equiv 1$. Dans ce cas, on a

$$\frac{1}{g(x_0+h)} - \frac{1}{g(x_0)} = -\frac{g(x_0+h) - g(x_0)}{g(x_0)g(x_0+h)},$$

et le résultat découle de 4.36.e, compte tenu de 7.12.

7.14. Exemples.

a. Une constante $f(x) \equiv c$ a, évidemment, 0 pour dérivée.

b. La fonction $f(x) \equiv x$ a 1 pour dérivée.

c. Compte tenu de 7.13, nous voyons que tout polynôme

$$P(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$$

ainsi que toute fonction rationnelle

$$\frac{P(x)}{Q(x)} = \frac{a_0x^n + a_1x^{n-1} + \dots + a_n}{b_0x^m + b_1x^{m-1} + \dots + b_m}$$

possède une dérivée au moins en tout point x_0 tel que $Q(x_0) \neq 0$.

Les dérivées des fonctions rationnelles peuvent être calculées en se servant des formules 7.13 (1)-(4) et de la formule

$$(x^n)' = nx^{n-1} \quad (1)$$

qui s'établit par récurrence à partir de 7.13 (3).

d. La dérivée d'un produit de n facteurs est calculée d'après la formule

$$[f_1(x) \dots f_n(x)]' = f_1'(x) f_2(x) \dots f_n(x) + f_1(x) f_2'(x) \dots \\ \dots f_n(x) + \dots + f_1(x) f_2(x) \dots f_n'(x),$$

ce qui se démontre par récurrence en se basant sur la formule 7.13 (3).

e. Trouvons la dérivée d'un déterminant formé de fonctions dérivables :

$$W(x) = \begin{vmatrix} u_{11}(x) & u_{12}(x) & \dots & u_{1n}(x) \\ u_{21}(x) & u_{22}(x) & \dots & u_{2n}(x) \\ \dots & \dots & \dots & \dots \\ u_{n1}(x) & u_{n2}(x) & \dots & u_{nn}(x) \end{vmatrix}.$$

Par définition, le déterminant $W(x)$ est la somme algébrique de $n!$ termes précédé chacun d'un signe déterminé et représentant un produit de n facteurs obtenus en prenant un élément et un seul dans chaque colonne et dans chaque ligne. En dérivant chaque terme selon la règle d et en réunissant d'abord les termes dans lesquels l'on a dérivé le facteur provenant de la première colonne, puis les termes dans lesquels l'on a dérivé le facteur de la deuxième

colonne, etc., on aboutit à la formule

$$\begin{aligned}
 W'(x) = & \begin{vmatrix} u'_{11}(x) & u_{12}(x) & \dots & u_{1n}(x) \\ u'_{21}(x) & u_{22}(x) & \dots & u_{2n}(x) \\ \dots & \dots & \dots & \dots \\ u'_{n1}(x) & u_{n2}(x) & \dots & u_{nn}(x) \end{vmatrix} + \\
 & + \begin{vmatrix} u_{11}(x) & u'_{12}(x) & \dots & u_{1n}(x) \\ u_{21}(x) & u'_{22}(x) & \dots & u_{2n}(x) \\ \dots & \dots & \dots & \dots \\ u_{n1}(x) & u'_{n2}(x) & \dots & u_{nn}(x) \end{vmatrix} + \dots + \\
 & + \begin{vmatrix} u_{11}(x) & u_{12}(x) & \dots & u'_{1n}(x) \\ u_{21}(x) & u_{22}(x) & \dots & u'_{2n}(x) \\ \dots & \dots & \dots & \dots \\ u_{n1}(x) & u_{n2}(x) & \dots & u'_{nn}(x) \end{vmatrix}.
 \end{aligned}$$

7.15.a. Dérivée d'une fonction composée.

Si la fonction $y = f(x)$ est dérivable pour $x = x_0$ et la fonction $z = g(y)$ est définie dans un intervalle contenant le point $y_0 = f(x_0)$ et dérivable pour $y = y_0$, alors la fonction composée $z = g[f(x)]$ est dérivable pour $x = x_0$, de sorte que

$$z'(x_0) = g'(y_0) f'(x_0). \quad (1)$$

Démonstration. Selon la définition de la dérivée,

$$y - y_0 = f(x) - f(x_0) = (x - x_0) [f'(x_0) + e(x)],$$

$$e(x) \rightarrow 0 \text{ pour } x \rightarrow x_0,$$

$$g(y) - g(y_0) = (y - y_0) [g'(y_0) + \delta(y)],$$

$$\delta(y) \rightarrow 0 \text{ pour } y \rightarrow y_0.$$

Donc

$$\begin{aligned}
 z(x) - z(x_0) &= g[f(x)] - g[f(x_0)] = g(y) - g(y_0) = \\
 &= (y - y_0) [g'(y_0) + \delta(y)] = \\
 &= (x - x_0) [f'(x_0) + e(x)] [g'(y_0) + \delta(y)].
 \end{aligned}$$

Lorsque $x \rightarrow x_0$, on a $e(x) \rightarrow 0$ (nous l'avons vu plus haut). Ensuite, en vertu de la continuité de la fonction $f(x)$ au point x_0 , on a aussi $y \rightarrow y_0$, donc $\delta(y) \rightarrow 0$. En passant à la limite dans l'égalité

$$\frac{z(x) - z(x_0)}{x - x_0} = [f'(x_0) + e(x)] [g'(y_0) + \delta(y)],$$

on obtient le résultat cherché de 4.36.b.

b. Exemple. La dérivée de la fonction

$$(1 + x^2)^{99}$$

peut être trouvée d'après les règles 7.14, après avoir développé le binôme selon la formule de Newton. Il est cependant plus simple de représenter la fonction donnée comme fonction composée :

$$y = u^{99}, \quad u = 1 + x^2.$$

Alors

$$y'(u) = 99u^{98}, \quad u'(x) = 2x$$

et, d'après (1),

$$y'(x) = 99u^{98} \cdot 2x = 99(1 + x^2)^{98} \cdot 2x.$$

7.16. Dérivée d'une fonction inverse. Soit $y = f(x)$ une fonction continue et croissante dans un intervalle (a, b) , et soit $x = \varphi(y)$ sa fonction inverse définie, au moins, dans un voisinage d'un point $y = y_0 = f(x_0)$, $x_0 \in (a, b)$. Alors, si la fonction $y = f(x)$ est dérivable pour $x = x_0$ et $f'(x_0) \neq 0$, la fonction $x = \varphi(y)$ est dérivable pour $y = y_0$ et

$$\varphi'(y_0) = \frac{1}{f'(x_0)}.$$

Démonstration. Considérons le rapport

$$\frac{\varphi(y) - \varphi(y_0)}{y - y_0}. \quad (1)$$

Par définition d'une fonction inverse,

$$\varphi(y) = x, \quad \varphi(y_0) = x_0, \quad y = f(x), \quad y_0 = f(x_0),$$

de sorte que le rapport (1) peut être mis sous la forme

$$\frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}}.$$

Une fonction et son inverse étant continues toutes les deux (5.38), il résulte de $y \rightarrow y_0$ que $x \rightarrow x_0$, et, $f(x)$ étant dérivable pour $x = x_0$ par hypothèse, on a

$$\frac{\varphi(y) - \varphi(y_0)}{y - y_0} \rightarrow \frac{1}{f'(x_0)},$$

ce qu'il fallait démontrer.

7.17. Dérivée du logarithme et des fonctions qui y sont associées.

a. Soit $y = \log_a x$. Si $x_0 > 0$, nous avons

$$\begin{aligned} \frac{\log_a(x_0+h) - \log_a x_0}{h} &= \log_a \left(\frac{x_0+h}{x_0} \right)^{\frac{1}{h}} = \\ &= \log_a \left(1 + \frac{h}{x_0} \right)^{\frac{1}{h}} = \frac{1}{h} \log_a \left(1 + \frac{h}{x_0} \right) = \\ &= \frac{1}{x_0} \cdot \frac{x_0}{h} \log_a \left(1 + \frac{h}{x_0} \right) \xrightarrow{(h \rightarrow 0)} \frac{\log_a e}{x_0} \end{aligned}$$

en vertu de 5.59.d. Donc

$$(\log_a x)'_{x=x_0} = \frac{\log_a e}{x_0}.$$

Cette formule prend une forme particulièrement simple pour $a = e$, lorsqu'il s'agit du logarithme naturel (5.59.d):

$$(\ln x)' = \frac{1}{x}. \quad (1)$$

b. Trouvons la dérivée de la fonction $y = \ln(-x)$ définie pour $x < 0$. D'après la règle de dérivation d'une fonction composée 7.15.a, nous obtenons

$$[\ln(-x)]' = \frac{1}{-x} \cdot (-x)' = -\frac{1}{x} \cdot (-1) = \frac{1}{x}. \quad (2)$$

On réunit les formules (1) et (2) en une seule valable pour tous les $x \neq 0$:

$$(\ln|x|)' = \frac{1}{x}. \quad (3)$$

c. L'exponentielle $x = a^y$ est l'inverse de la fonction $y = \log_a x$; donc, d'après 7.16, nous avons:

$$(a^y)'_{y=y_0} = \frac{1}{\log_a' x_0} = \frac{x_0}{\log_a e} = \frac{a^{y_0}}{\log_a e} = a^{y_0} \log_e a = a^{y_0} \ln a.$$

En remplaçant y_0 par x , $-\infty < x < \infty$, on obtient

$$(a^x)' = a^x \ln a. \quad (4)$$

Cette formule devient particulièrement simple pour $a = e$:

$$(e^x)' = e^x. \quad (5)$$

d. Dérivée de la fonction puissance d'exposant quelconque. La fonction

$$y = x^\alpha$$

définie pour $x > 0$ (5.54) peut être mise sous la forme

$$y = e^{\alpha \ln x}.$$

Il en résulte qu'elle possède une dérivée que l'on peut trouver suivant la règle 7.15 et compte tenu des résultats a-c. Nous avons

$$y'(x) = e^{\alpha \ln x} \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1}. \quad (6)$$

Dans le cas de α naturel, on obtient la formule 7.14 (1) (on devait bien s'y attendre).

e. La même idée peut aider à trouver la dérivée d'une fonction plus compliquée

$$y = f(x)^{g(x)} = e^{g(x) \ln f(x)},$$

ce que nous laissons au lecteur.

7.18. Dérivée des fonctions trigonométriques et de leurs inverses.

a. Pour la fonction $y = \sin x$, nous avons, d'après 5.63:

$$\sin(x+h) - \sin x = 2 \sin \frac{h}{2} \cdot \cos \left(x + \frac{h}{2}\right),$$

d'où

$$\frac{\sin(x+h) - \sin x}{h} = \frac{\sin \frac{h}{2}}{\frac{h}{2}} \cdot \cos \left(x + \frac{h}{2}\right).$$

Etant donnés 5.64.c et la continuité de la fonction $\cos x$, on trouve

$$(\sin x)' = \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} = \cos x. \quad (1)$$

b. Ensuite, en appliquant les formules 5.65 (2), (3) on obtient

$$(\cos x)' = \left[\sin \left(\frac{\pi}{2} + x \right) \right]' = \cos \left(\frac{\pi}{2} + x \right) = -\sin x. \quad (2)$$

c. Enfin, d'après la règle 7.13 (4) pour $x \neq \frac{2k+1}{2} \pi$, $k=0, \pm 1, \pm 2, \dots$, on a

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x} \right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}. \quad (3)$$

d. Si $x = \arcsin u$, $u = \sin x$, alors, suivant la règle 7.16, pour $-\frac{\pi}{2} < x < \frac{\pi}{2}$ et $-1 < u < 1$, on a

$$(\arcsin u)' = \frac{1}{(\sin x)'} = \frac{1}{\cos x} = \frac{1}{+\sqrt{1-\sin^2 x}} = \frac{1}{+\sqrt{1-u^2}}. \quad (4)$$

e. D'une façon analogue, pour $-\pi < x < 0$ et $-1 < u < 1$, on a

$$(\arccos u)' = \frac{1}{(\cos x)'} = -\frac{1}{\sin x} = -\frac{1}{-\sqrt{1-\cos^2 x}} = +\frac{1}{\sqrt{1-u^2}}. \quad (5)$$

On aurait pu obtenir l'expression (5) à partir de (4) en se servant de l'égalité 5.67 (1)

$$\arccos u + \frac{\pi}{2} = \arcsin u.$$

D'une façon plus exacte, l'égalité (5) est valable pour la fonction croissante $\arccos u$ que nous avons désignée, dans 5.67, par $\arccos_c u$. En ce qui concerne la fonction décroissante $\arccos_d u$, l'égalité (5) doit être remplacée par

$$(\arccos_d u)' = -\frac{1}{\sqrt{1-u^2}}.$$

f. Enfin, pour $-\frac{\pi}{2} < x < \frac{\pi}{2}$, $-\infty < u < \infty$, on a

$$(\arctg u)' = \frac{1}{(\lg x)'} = \cos^2 x = \frac{1}{1 + \lg^2 x} = \frac{1}{1 + u^2}. \quad (6)$$

Nous laissons au lecteur le soin de trouver les dérivées d'autres fonctions trigonométriques et de leurs inverses.

7.19. Dérivées à gauche et à droite.

a. Soit une fonction $f(x)$ définie dans un intervalle $x_0 \leq x < b$. Nous dirons qu'elle possède une dérivée à droite pour $x = x_0$ s'il existe

$$f'_{dr}(x_0) = \lim_{h \searrow 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

La quantité $f'_{dr}(x_0)$ est aussi désignée par $f'(x_0 + 0)$.

b. D'une façon analogue, nous dirons qu'une fonction $f(x)$ définie dans un intervalle $a < x \leq x_0$ possède, pour $x = x_0$, une dérivée à gauche s'il existe

$$f'_g(x_0) = \lim_{h \nearrow 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

La quantité $f'_g(x_0)$ est aussi désignée par $f'(x_0 - 0)$.

c. Si une fonction $f(x)$ est définie dans un intervalle (a, b) contenant un point x_0 , on peut discuter l'existence de ses dérivées à gauche et à droite au point x_0 . Lorsque la dérivée $f'(x_0)$ existe,

alors il existe, bien entendu, les dérivées $f'_{dr}(x_0)$ et $f'_g(x_0)$, et, de plus, $f'_{dr}(x_0) = f'_g(x_0) = f'(x_0)$. Mais il peut arriver que $f'_{dr}(x_0)$ et $f'_g(x_0)$ existent et $f'(x_0)$ non (fig. 7.2). Lorsqu'il existe $f'_{dr}(x_0)$ et $f'_g(x_0)$ et que l'égalité $f'_{dr}(x_0) = f'_g(x_0)$ est valable, $f'(x_0)$ existe aussi, ce qui est facile à prouver (cf. 4.16.c).

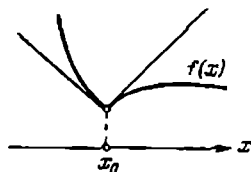


Fig. 7.2.

d. La demi-droite définie par l'équation

$$y = f'(x_0) + f_{dr}'(x_0)(x - x_0) \quad (x \geq x_0)$$

s'appelle *demi-tangente à droite* à la courbe $y = f(x)$ pour $x = x_0$. D'une façon analogue, la demi-droite définie par l'équation

$$y = f(x_0) + f_{g}'(x_0)(x - x_0) \quad (x \leq x_0)$$

s'appelle *demi-tangente à gauche* à la courbe $y = f(x)$ pour $x = x_0$.

e. La proposition suivante généralise le théorème 7.16 au cas des dérivées à gauche et à droite :

Soient $y = f(x)$ une fonction continue et croissante dans un intervalle (a, b) et $x = \varphi(y)$ sa fonction inverse. Si la fonction $y = f(x)$ possède, pour un $x = x_0$, une dérivée à droite $f'(x_0 + 0) \neq 0$, alors la fonction $\varphi(y)$ possède, pour $y_0 = f(x_0)$, une dérivée à droite $\varphi'(y_0 + 0)$ telle que

$$\varphi'(y_0 + 0) = \frac{1}{f'(x_0 + 0)}.$$

La démonstration est analogue à celle de 7.16, en considérant seules les valeurs $x \geq x_0$ et $y \geq y_0$.

§ 7.2. Deuxième définition de la dérivée

7.21. A présent, nous allons analyser des propriétés générales des fonctions dérivables. Avant tout, considérons une autre manière d'introduire la notion de dérivée qui est importante pour ce qui suit. Construisons les fonctions linéaires $Ax + B$ ayant, pour $x = x_0$, la même valeur qu'une fonction donnée $y = f(x)$; on peut les décrire par l'équation

$$y_A = A(x - x_0) + f(x_0) = A \cdot h + f(x_0) \quad (h = x - x_0). \quad (1)$$

Deux fonctions de ce genre $y_A = A(x - x_0) + f(x_0)$ et $y_B = B(x - x_0) + f(x_0)$ diffèrent par une quantité proportionnelle à $x - x_0$:

$$y_A - y_B = (A - B)(x - x_0) = (A - B)h.$$

Choisissons, parmi les fonctions (1), l'une qui diffère de la fonction $y = f(x)$ par un infiniment petit par rapport à $h = x - x_0$ (4.38), c'est-à-dire telle que

$$y - y_A = \varepsilon(h) \cdot h, \quad (2)$$

où $\varepsilon(h)$ tend vers 0 lorsque $h \rightarrow 0$.

Supposons qu'un tel A soit trouvé. Alors il résulte de (2) que

$$\frac{f(x) - A(x - x_0) - f(x_0)}{h} = \frac{f(x_0 + h) - f(x_0)}{h} - A = \varepsilon(h) \rightarrow 0,$$

de sorte que le rapport

$$\frac{f(x_0+h)-f(x_0)}{h}$$

possède une limite A lorsque $h \rightarrow 0$. Donc, si une fonction linéaire cherchée existe, alors la fonction $f(x)$ est dérivable pour $x = x_0$, et le coefficient cherché A vaut $f'(x_0)$.

7.22. Réciproquement, si la fonction $f(x)$ est dérivable pour $x = x_0$, nous avons

$$\frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) = \varepsilon(h) \rightarrow 0,$$

d'où

$$f(x_0+h) - [f'(x_0)h + f(x_0)] = \varepsilon(h) \cdot h, \quad (1)$$

de sorte que la fonction linéaire

$$Y = f'(x_0)h + f(x_0)$$

de la variable $h = x - x_0$ satisfait à la condition imposée: son écart de la fonction $f(x)$ est un infiniment petit par rapport à $h = x - x_0$.

La droite dans le plan (x, y) définie par l'équation

$$Y = f'(x_0)(x - x_0) + f(x_0)$$

est, d'après la définition 7.11, la tangente à la courbe $y = f(x)$ pour $x = x_0$. Quel que soit $\varepsilon > 0$, il existe un $\delta > 0$ tel que $|\varepsilon(h)| < \varepsilon$ dès que $|h| < \delta$; donc, pour ces h , on a

$$-eh \leq f(x) - f'(x_0)h - f(x_0) \leq eh$$

ou

$$f(x_0) + [f'(x_0) - \varepsilon]h \leq f(x) \leq f(x_0) + [f'(x_0) + \varepsilon]h. \quad (2)$$

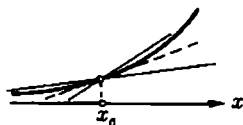


Fig. 7.3.

Voici l'interprétation géométrique de l'inégalité (2): le graphique d'une fonction dérivable en un point $x = x_0$ est compris, pour x assez proche de x_0 , entre deux droites dont chacune forme avec la tangente un angle aussi petit que l'on veut (fig. 7.3).

7.23. En tant que conséquence nous avons:

a. Si $f'(x_0) > 0$, il existe un $\delta > 0$ tel que, pour $0 < h < \delta$,

$$f(x_0 - h) < f(x_0) < f(x_0 + h). \quad (1)$$

b. Et si $f'(x_0) < 0$, alors, pour $0 < h < \delta$,

$$f(x_0 - h) > f(x_0) > f(x_0 + h). \quad (2)$$

7.24. Cette propriété est à la base d'une règle de calcul des points d'extrémum local.

D é f i n i t i o n. Nous dirons que la fonction $f(x)$ a un maximum local en un point $c \in (a, b)$ s'il existe un $h > 0$ tel que, pour tous les $x \in (c - h, c + h)$,

$$f(x) \leq f(c)$$

(fig. 7.4). D'une façon analogue, nous dirons que la fonction $f(x)$ a un minimum local en un point $c \in (a, b)$ s'il existe un $h > 0$ tel que, pour tous les $x \in (c - h, c + h)$,

$$f(x) \geq f(c)$$

(fig. 7.5). Les points de minimum ou maximum local sont appelés points d'extrémum local. Si la fonction $f(x)$ est dérivable pour $x = c$

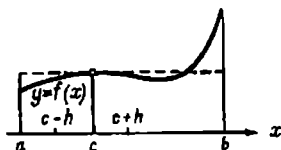


Fig. 7.4.

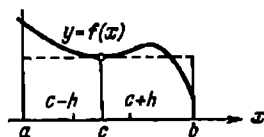


Fig. 7.5.

et que $f'(c) \neq 0$, alors les inégalités 7.23 (1) et (2) montrent que le point c ne peut être un point d'extrémum local. Il en résulte qu'en tout point d'extrémum local d'une fonction dérivable $f(x)$, l'égalité

$$f'(c) = 0 \quad (1)$$

est satisfaite.

Donc, pour trouver les points d'extrémum local d'une fonction (partout) dérivable, il faut analyser l'équation (1). Les points cherchés sont parmi ses racines. Mais il se peut que $f'(c) = 0$ sans que l'extrémum local existe. Ceci a lieu, par exemple, pour la fonction $f(x) = x^3$ avec $c = 0$. Une analyse plus détaillée d'extréma est exposée plus loin (§ 7.5 et § 8.4).

§ 7.3. Différentielle

7.31. On voit de l'égalité 7.22 (1) que l'accroissement de la fonction $f(x)$, lorsque la variable indépendante x varie de $x = x_0$ à $x = x_0 + h$, est composé de deux parties: la première, $f'(x_0)h$, est linéaire par rapport au déplacement h de la variable x , et la seconde, $e(h) \cdot h$, est infiniment petite par rapport à h . Géométriquement, la première partie représente l'accroissement mesuré jusqu'à

la tangente, la seconde l'accroissement à partir de la tangente jusqu'à la valeur même $f(x_0 + h)$ (fig. 7.6). Le terme $f'(x_0)h$ s'appelle *partie linéaire principale de l'accroissement de la fonction*. Donc, l'existence de la dérivée implique la possibilité de dégager la partie linéaire principale de l'accroissement.

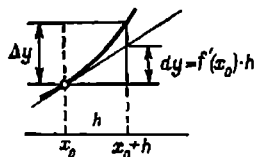


Fig. 7.6.

Inversement, le raisonnement fait montrer que la possibilité de dégager la partie linéaire principale de l'accroissement de la fonction garantit l'existence de la dérivée.

7.32. Différentielle. Supposons que $f(x)$ soit dérivable pour $x = x_0$. La quantité $h = x - x_0$ est encore désignée par dx , $f'(x_0)h = f'(x_0)dx$ par $dy(x_0)$ ou dy . Le nombre dy s'appelle *différentielle de la fonction $f(x)$ au point $x = x_0$ pour la différentielle dx de la variable indépendante*. Ainsi, dy est une fonction linéaire de dx .

Etant donnés dx et dy , on peut mettre $f'(x_0)$ sous la forme de rapport des différentielles:

$$f'(x_0) = \frac{dy}{dx}.$$

7.33. Les règles de calcul des dérivées que l'on a décrites dans § 7.1 amènent aux règles correspondantes de calcul des différentielles. Notamment, en multipliant par dx les formules 7.13 (1)-(4), on trouve

$$d[f + g] = df + dg; \quad (1)$$

$$d(\alpha f) = \alpha df; \quad (2)$$

$$d[f \cdot g] = f dg + df \cdot g; \quad (3)$$

$$d\left[\frac{f}{g}\right] = \frac{g \cdot df - dg \cdot f}{g^2(x_0)}; \quad (4)$$

la dernière égalité a lieu, comme dans 7.13 (4), à condition que $g(x_0) \neq 0$.

7.34. Différentielle d'une fonction composée. Après avoir multiplié la formule 7.15 (1) par dx , on a

$$dz = g'(y_0) y'(x_0) dx = g'(y_0) dy.$$

Or, si y était une variable indépendante et non pas une fonction de x , alors, par définition, la différentielle de la fonction z serait de la même forme:

$$dz = g'(y_0) dy.$$

Donc, la différentielle d'une fonction a la même forme que l'argument de la fonction soit une variable indépendante ou une fonction d'un autre argument.

On utilisera ce fait pour dériver les fonctions composées. Par exemple, la dérivée de $(x^3 + 1)^{99}$ (7.15.b) peut être trouvée à l'aide des différentielles :

$d(x^3 + 1)^{99} = 99(x^3 + 1)^{98} d(x^3 + 1) = 99(x^3 + 1)^{98} \cdot 2x dx$,
d'où

$$((x^3 + 1)^{99})' = \frac{d(x^3 + 1)^{99}}{dx} = 99(x^3 + 1)^{98} \cdot 2x.$$

§ 7.4. Théorèmes des accroissements finis

7.41. Théorème de Rolle. *Si une fonction (finie) $f(x)$ est continue sur un intervalle $[a, b]$ (qui peut être infini), de sorte que $f(a) = f(b)$, et si elle est dérivable en tout point de l'intervalle (a, b) , alors il existe un point $c \in (a, b)$ pour lequel $f'(c) = 0$.*

Démonstration. En vertu de 5.16.c, il existe un point $c \in (a, b)$ en lequel l'égalité $f(c) = \sup\{f(x)\}$ (ou bien $f(c) = \inf\{f(x)\}$) est satisfaite. Donc, c est un point d'extrémum local de la fonction $f(x)$. D'après 7.24, on a $f'(c) = 0$, ce qu'il fallait démontrer.

7.42. Conséquence. *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$ et dérivable dans l'intervalle (a, b) , et si l'inégalité $f'(x) \neq 0$ a lieu partout dans (a, b) , alors $f(b) \neq f(a)$.*

Ceci se démontre par l'absurde en appliquant 7.41.

7.43. Théorème de Cauchy. *Si deux fonctions (finies) $f(x)$ et $g(x)$ sont continues sur un intervalle $a \leq x \leq b$ (qui peut être infini) et dérivables en tout point intérieur de cet intervalle, de sorte que $g'(x)$ ne s'y annule pas, alors il existe un point $c \in (a, b)$ pour lequel*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}. \quad (1)$$

Démonstration. Notons que $g(b) \neq g(a)$ en vertu de 7.42. La fonction $\varphi(x) = f(x) - Ag(x)$, avec n'importe quelle constante A , est continue sur l'intervalle $[a, b]$ et dérivable en tout son point intérieur puisque les fonctions $f(x)$ et $g(x)$ le sont. Choisissons un A de façon que la fonction $\varphi(x)$ vérifie l'hypothèse $\varphi(b) = \varphi(a)$ du théorème de Rolle. Nous avons alors l'équation

$$f(a) - Ag(a) = f(b) - Ag(b),$$

d'où

$$A = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Appliquons le théorème de Rolle : il existe un point $c \in (a, b)$ en lequel

$$\varphi'(c) = f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(c) = 0,$$

d'où (1).

7.44. Conséquence (théorème de Lagrange). Si une fonction $f(x)$ est continue sur un intervalle fini $[a, b]$ et dérivable en tout point intérieur, alors il existe un point $c \in (a, b)$ tel que

$$f'(c) = \frac{f(b)-f(a)}{b-a}. \quad (1)$$

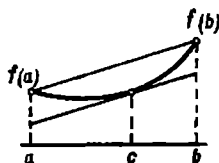


Fig. 7.7.

Pour le démontrer, il faut poser $g(x) = x$ dans 7.43. La formule (1) se met alors sous la forme

$$f(b) = f(a) + f'(c)(b-a); \quad (2)$$

c'est la *formule des accroissements finis*.

Géométriquement, c est un tel point du segment $[a, b]$ que la tangente à la courbe $y = f(x)$ au point $(c, f(c))$ est parallèle à la corde qui joint les points $(a, f(a))$ et $(b, f(b))$ (fig. 7.7).

7.45. Conséquences.

a. Si $f'(x) > 0$ partout dans (a, b) , alors la fonction $f(x)$ croît sur $[a, b]$.

b. Si $f'(x) < 0$ partout dans (a, b) , alors la fonction $f(x)$ décroît sur $[a, b]$.

c. Si $f'(x) = 0$, alors $f(x)$ est constante.

En abordant la démonstration, remplaçons, dans 7.44 (1), a par x' , b par x'' , où $[x', x''] \subset [a, b]$.

Ensuite,

(a) si $f'(x) > 0$, nous avons

$$f(x'') = f(x') + (x'' - x') f'(c) > f(x');$$

(b) si $f'(x) < 0$, alors

$$f(x'') = f(x') + (x'' - x') f'(c) < f(x');$$

(c) et si $f'(x) = 0$, alors

$$f(x'') = f(x') + (x'' - x') f'(c) = f(x'),$$

ce qui démontre la proposition.

d. Exemple. Soit $f(x) = \sin x$, $f'(x) = \cos x$. Dans les intervalles $-\frac{\pi}{2} + 2k\pi < x < \frac{\pi}{2} + 2k\pi$, la fonction $\cos x$ est positive (5.65), et nous concluons que la fonction $\sin x$ croît dans ces intervalles; dans les intervalles $\frac{\pi}{2} + 2k\pi < x < \frac{3\pi}{2} + 2k\pi$, la fonction $\cos x$ est négative, donc la fonction $\sin x$ décroît. D'une

façon analogue, si l'on pose $f(x) = \cos x$, $f'(x) = -\sin x$, on verra que la fonction $\cos x$ croît dans les intervalles $(2k-1)\pi < x < 2k\pi$ et décroît dans les intervalles $2k\pi < x < (2k+1)\pi$. (A proprement parler, nous avons fait de pareilles conclusions dans 5.65, mais les raisonnements présents sont plus succincts.)

§ 7.5. Position d'une courbe par rapport à sa tangente

7.51. Supposons qu'une fonction $y = f(x)$ soit continue dans un intervalle (a, b) contenant un point x_0 et qu'elle y possède la dérivée $f'(x)$. Nous considérons ici la position relative du graphique de la fonction $y = f(x)$ et de sa tangente au point $x = x_0$:

$$Y(x) = f(x_0) + f'(x_0) \cdot (x - x_0).$$

Introduisons les définitions suivantes. Un point x_0 s'appelle *point de convexité dirigée vers le haut* de la fonction $f(x)$ si, au voisinage de x_0 (sauf en ce point même), on a l'inégalité $f(x) < Y(x)$, ou, autrement dit, si le graphique de la fonction $f(x)$, au voisinage du point x_0 , est au-dessous de la tangente au point $x = x_0$. D'une façon analogue, un point x_0 s'appelle *point de convexité dirigée vers le bas* de la fonction $f(x)$ si, au voisinage de x_0 , on a l'inégalité

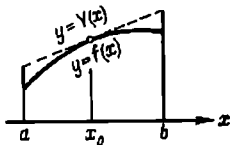


Fig. 7.8.

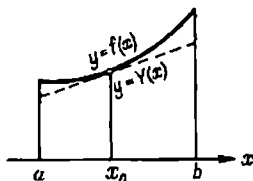


Fig. 7.9.

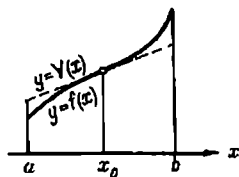


Fig. 7.10.

$f(x) > Y(x)$, ou, ce qui revient au même, si le graphique de la fonction est au-dessus de la tangente. Si tout point $x \in (a, b)$ est un point de convexité dirigée vers le haut (vers le bas) pour la fonction $y = f(x)$, alors la fonction $f(x)$ est dite *convexe vers le haut* (vers le bas) dans l'intervalle (a, b) . Un point x_0 s'appelle *point d'inflexion* de la fonction $f(x)$ si la courbe $f(x)$ est disposée, pour $x < x_0$, d'un côté de la tangente au point $x = x_0$, et, pour $x > x_0$, de l'autre. Les fig. 7.8-7.10 illustrent les cas mentionnés.

Ci-dessous on donne des conditions suffisantes (analytiques) pour que tel ou tel cas se produise.

7.52.a. Si l'inégalité $f'(\eta) < f'(x_0)$ a lieu pour tous les $\eta \in (x_0, b)$, alors la courbe $f(x)$ est, pour $x > x_0$, au-dessous de la tangente au point $x = x_0$.

b. Si l'inégalité $f'(\eta) > f'(x_0)$ a lieu pour tous les $\eta \in (x_0, b)$, alors la courbe $f(x)$ est, pour $x > x_0$, au-dessus de cette tangente.

En effet, dans la première hypothèse, nous avons pour $x > x_0$, d'après la formule de Lagrange :

$f(x) = f(x_0) + f'(\eta)(x - x_0) < f(x_0) + f'(x_0)(x - x_0) = Y(x)$, ce qu'il fallait démontrer ; pour la deuxième hypothèse, le raisonnement est analogue.

7.53. De même, si l'inégalité $f'(\xi) > f'(x_0)$ a lieu pour tous les $\xi \in (a, x_0)$, alors la courbe $f(x)$ est, pour $x < x_0$, au-dessous de la tangente au point $x = x_0$; si l'inégalité $f'(\xi) < f'(x_0)$ a lieu pour tous les $\xi \in (a, x_0)$, alors la courbe $f(x)$ est, pour $x < x_0$, au-dessus de cette tangente.

Pour le démontrer on applique, tout comme en 7.52, la formule de Lagrange à l'intervalle $[x, x_0]$.

7.54. En combinant les cas décrits dans 7.52 et 7.53 on aboutit aux résultats suivants :

Si, pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$, on a l'inégalité

$$f'(\xi) > f'(x_0) > f'(\eta),$$

alors x_0 est un point de convexité dirigée vers le haut pour la fonction $f(x)$; si, pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$, on a l'inégalité

$$f'(\xi) < f'(x_0) < f'(\eta),$$

alors x_0 est un point de convexité dirigée vers le bas pour la fonction $f(x)$. Enfin, si pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$ ont lieu les inégalités

$$f'(\xi) > f'(x_0), \quad f'(\eta) > f'(x_0),$$

ou bien les inégalités

$$f'(\xi) < f'(x_0), \quad f'(\eta) < f'(x_0),$$

alors x_0 est un point d'inflexion de la fonction $f(x)$.

7.55. La conclusion suivante est immédiate :

Si la fonction $f'(x)$ croît pour $a < x < b$, alors la fonction $f(x)$ est convexe vers le bas dans l'intervalle (a, b) ; si $f'(x)$ décroît pour $a < x < b$, la fonction $f(x)$ est convexe vers le haut dans l'intervalle (a, b) .

E x e m p l e. Posons $f(x) = \sin x$, $f'(x) = \cos x$. La fonction $\cos x$ étant croissante dans les intervalles $(2k-1)\pi < x < 2k\pi$ (7.45), la fonction $\sin x$ y est convexe vers le bas; la fonction $\cos x$ étant décroissante dans les intervalles $2k\pi < x < (2k+1)\pi$, la fonction $\sin x$ y est convexe vers le haut. D'une façon analogue, la fonction $\cos x$ est convexe vers le bas dans les intervalles $(2k + \frac{1}{2})\pi < x < (2k + \frac{3}{2})\pi$ et vers le haut dans les intervalles $(2k - \frac{1}{2})\pi < x < (2k + \frac{1}{2})\pi$. (Ceci explique la forme des graphiques des fonctions $\sin x$ et $\cos x$ que l'on a donnés dans 5.65.)

Dans 8.32-8.33, on formulera des conditions analytiques de détermination des points de convexité et d'inflexion qui seront basées sur des propriétés de dérivées d'ordre supérieur de la fonction $f(x)$.

7.56. Si $f'(x_0) = 0$, alors $Y(x) \equiv f(x_0)$, et la tangente au point $x = x_0$ est horizontale. Si, de plus, le point x_0 est un point de convexité vers le bas, i.e. la courbe $f(x)$ est au-dessus de la tangente, alors nous avons $f(x) > f(x_0)$ au voisinage du point x_0 , donc, x_0 est un point de minimum local de la fonction $f(x)$. De même, si $f'(x_0) = 0$ et le point x_0 est un point de convexité dirigée vers le haut, alors il est un point de maximum local. Nous tirons maintenant de 7.54 les conditions suffisantes d'existence d'un minimum ou d'un maximum :

T h é o r è m e Soit $f'(x_0) = 0$ Si, de plus, les inégalités

$$f'(\xi) < 0, \quad f'(\eta) > 0$$

ont lieu pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$, alors x_0 est un point de minimum local de la fonction $f(x)$. Si, au contraire, on a, pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$, les inégalités

$$f'(\xi) > 0, \quad f'(\eta) < 0,$$

alors x_0 est un point de maximum local de la fonction $f(x)$.

7.57. Soit, comme dans 7.56, $f'(x_0) = 0$, et les inégalités

$$f'(\xi) < 0, \quad f'(\eta) < 0$$

ont lieu pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$. Alors, d'après 7.53, le point x_0 est un point d'inflexion de la courbe $y = f(x)$ qui passe, au point x_0 , d'un côté de la tangente à l'autre, de sorte que le point x_0 n'est pas un point de minimum ni un point de maximum. Une situation analogue a lieu lorsque, pour tous les $\xi \in (a, x_0)$ et $\eta \in (x_0, b)$, on a les inégalités

$$f'(\xi) > 0, \quad f'(\eta) > 0.$$

§ 7.6. Règles de L'Hospital

On se sert souvent des théorèmes de ce paragraphe pour calculer des limites.

7.61. Première règle de L'Hospital. Soient $f(x)$ et $g(x)$ des fonctions (finies) continues sur un segment $[a, b]$ (qui peut être infini) et dérivables dans l'intervalle (a, b) , et soit $g'(x) \neq 0$ partout dans (a, b) . Soit, enfin,

$$g(a) = f(a) = 0.$$

On dit alors que le rapport $\frac{f(x)}{g(x)}$ représente, pour $x \searrow a$, une *indétermination de la forme* $\frac{0}{0}$.

Théorème. Si, dans les conditions formulées ci-dessus,

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = A$$

(sur la droite numérique achevée), alors

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = A.$$

Démonstration. Supposons d'abord que $-\infty < A < \infty$. Pour un $\varepsilon > 0$ donné, choisissons un x_0 de façon à avoir, dans l'intervalle (a, x_0) , l'inégalité

$$\left| \frac{f'(x)}{g'(x)} - A \right| < \varepsilon. \quad (1)$$

Appliquons le théorème de Cauchy 7.43 à l'intervalle $[a, x_0]$, où $a < x < x_0$; en vertu de ce théorème, il existe un point $c \in (a, x_0)$ tel que

$$\frac{f(x)}{g(x)} = \frac{f(x_0) - f(a)}{g(x_0) - g(a)} = \frac{f'(c)}{g'(c)},$$

d'où, pour tous les x , $a < x < x_0$,

$$\left| \frac{f(x)}{g(x)} - A \right| = \left| \frac{f'(c)}{g'(c)} - A \right| < \varepsilon.$$

Or, ceci veut dire que $A = \lim_{x \searrow a} \frac{f(x)}{g(x)}$.

Dans le cas où A est infini l'inégalité (1) est à remplacer par $\frac{f'(x)}{g'(x)} > \frac{1}{\varepsilon}$ ou $\frac{f'(x)}{g'(x)} < -\frac{1}{\varepsilon}$ selon le signe de A , pour le reste la démonstration ne varie pas.

7.62. Deuxième règle de L'Hospital. Soient $f(x)$ et $g(x)$ des fonctions continues et dérivables dans un intervalle (a, b) (peut-être infini), $g'(x)$ ne s'annulant pas dans (a, b) . Soit encore $g(a) = f(a) = \infty$; on dit alors que le rapport $\frac{f(x)}{g(x)}$ représente, pour $x \searrow a$, une *indétermination de la forme* $\frac{\infty}{\infty}$.

Théorème. Si, dans les conditions formulées ci-dessus,

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = A$$

(sur la droite numérique achevée), alors

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = A.$$

Démonstration. Soit d'abord A finl. Pour un $\varepsilon > 0$ donné, choisissons un x_0 de façon à avoir, dans l'intervalle (a, x_0) , l'inégalité

$$\left| \frac{f'(x)}{g'(x)} - A \right| < \varepsilon. \quad (1)$$

Définissons la fonction $D(x, x_0)$ par la condition

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} D(x, x_0).$$

Nous avons

$$D(x, x_0) = \frac{\frac{g(x) - g(x_0)}{g(x)}}{\frac{f(x) - f(x_0)}{f(x)}} = \frac{1 - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} \rightarrow 1 \quad (2)$$

pour $x \searrow a$. En appliquant le théorème de Cauchy à l'intervalle $[x, x_0]$ nous voyons que, pour un point $c \in [x, x_0]$,

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)} \cdot D(x, x_0) = \frac{f'(c)}{g'(c)} + \frac{f'(c)}{g'(c)} [D(x, x_0) - 1].$$

D'où, pour x tels que $|D(x, x_0) - 1| < \varepsilon$, nous trouvons

$$\begin{aligned} \left| \frac{f(x)}{g(x)} - A \right| &\leq \left| \frac{f'(c)}{g'(c)} - A \right| + \left| \frac{f'(c)}{g'(c)} \right| |D(x, x_0) - 1| \leq \\ &\leq \varepsilon + (|A| + \varepsilon) \varepsilon. \end{aligned} \quad (3)$$

Comme ε est arbitrairement petit, $\lim_{x \searrow a} \frac{f(x)}{g(x)} = A$, ce qu'il fallait démontrer

Dans le cas où $A = \infty$, l'inégalité (1) est remplacée par $\frac{f'(x)}{g'(x)} > \frac{1}{\varepsilon}$, et l'inégalité (3) par

$$\frac{f(x)}{g(x)} \geq \frac{f'(c)}{g'(c)} \cdot \frac{1}{2}$$

qui a lieu pour les x assez proches de A , en vertu de (2). Dans le cas où $A = -\infty$, les raisonnements sont analogues.

Exercices

1. Supposons que $f(x)$ soit définie pour $a \leq x \leq b$ et que, quels que soient x_1 et x_2 de $[a, b]$, elle satisfasse à l'inégalité

$$|f(x_1) - f(x_2)| \leq C |x_1 - x_2|^{1+\alpha}, \quad \alpha > 0.$$

Démontrer que $f(x)$ est constante.

2. Supposons que $f(x)$ soit définie et dérivable pour $x > c$. Soit, ensuite, $\lim_{x \rightarrow \infty} f'(x) = 0$. Alors, pour tout $h > 0$,

$$\lim_{x \rightarrow \infty} |f(x+h) - f(x)| = 0.$$

3. Une fonction $f(x)$ qui possède, sur un intervalle $[a, b]$, une dérivée continue, est uniformément dérivable sur $[a, b]$; en d'autres termes, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $|x_1 - x_2| < \delta$, $x_1, x_2 \in [a, b]$ implique l'inégalité

$$\left| f'(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} \right| < \varepsilon.$$

4. Montrer que la fonction $y = x^3 \sin \frac{1}{x}$ est partout dérivable sur $[0, 1]$, mais sa dérivée n'est pas continue.

5. La dérivée d'une fonction $f(x)$ partout dérivable sur $[a, b]$ prend toute valeur comprise entre $f'(a)$ et $f'(b)$ (théorème de Darboux).

6. Si une fonction $f(x)$ est convexe vers le bas sur (a, b) (7.51), alors, pour tout $x \in (a, b)$, on a l'inégalité

$$f(x) \leq Y(x) = \frac{f(b)(x-a) + f(a)(b-x)}{b-a}.$$

c'est-à-dire que le graphique de la fonction $f(x)$ n'a pas de points au-dessus de la corde qui joint les points d'abscisses a et b .

7. Le résultat de l'exercice 6 nous mène à une définition plus générale de la convexité d'une fonction que celle de 7.51 et qui n'utilise pas la dérivabilité: une fonction $y = f(x)$ s'appelle *convexe (vers le bas)* sur $[a, b]$ lorsque, sur tout intervalle $(\alpha, \beta) \subset [a, b]$, son graphique n'a aucun point au-dessus de la corde joignant les points d'abscisses α et β , c'est-à-dire que, pour tout $x \in (\alpha, \beta)$,

$$f(x) \leq \frac{f(\beta)(x-\alpha) + f(\alpha)(\beta-x)}{\beta-\alpha}.$$

Démontrer que, quels que soient les points $x_1, \dots, x_p$ de l'intervalle $[a, b]$

et les nombres $\lambda_1, \dots, \lambda_p$ tels que $0 \leq \lambda_i \leq 1$ ($i=1, \dots, p$) et $\sum_{j=1}^p \lambda_j = 1$,

on a

$$f\left(\sum_{j=1}^p \lambda_j x_j\right) \leq \sum_{j=1}^p \lambda_j f(x_j).$$

8. (Suite. Dans les exercices 8-13, on se sert de la définition de la convexité qu'on a introduite dans l'exercice 7.) Montrer qu'une fonction $f(x)$ est convexe sur un intervalle $[a, b]$ si, et seulement si, la pente

$$p(\alpha, x) = \frac{f(x) - f(\alpha)}{x - \alpha}$$

de la corde, pour $x > \alpha \geq a$, est une fonction non décroissante de x , quel que soit α fixe.

9. (Suite.) Une fonction $f(x)$ convexe sur un intervalle $[a, b]$ est continue en tout point $x \in (a, b)$ et possède ses dérivées finies à gauche et à droite, de sorte que

$$f'_g(x) \leq f'_{dr}(x).$$

10. (Suite.) Si $f(x)$ est convexe sur $[a, b]$, alors les fonctions $f'_g(x)$ et $f'_{dr}(x)$ ne sont pas décroissantes, et, pour n'importe quels α et β , $a < \alpha < \beta < b$, on a

$$f'_d(\alpha) \leq \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leq f'_g(\beta).$$

11. (Suite.) Toute fonction $f(x)$ convexe sur $[a, b]$ possède sa dérivée partout sur $[a, b]$, sauf, peut-être, en un ensemble dénombrable de points.

12. Supposons que $f(x)$ soit dérivable sur $[a, b]$ et que, pour deux points quelconques α, β où $a \leq \alpha < \beta \leq b$, il existe un seul point γ tel que

$$f'(\gamma) = \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

Montrer que ou bien $f(x)$ ou bien $-f(x)$ est convexe sur $[a, b]$.

13. Si une fonction $y = f(x)$ est convexe vers le bas au voisinage d'un point $x_0 \in (a, b)$, alors son graphique n'a pas de points au-dessous des demi-tangentes au point x_0 , c'est-à-dire que

$$f(x) \geq (x - x_0) f'_{dr}(x_0) + f(x_0) \quad (x > x_0),$$

$$f(x) \geq (x - x_0) f'_g(x_0) + f(x_0) \quad (x < x_0).$$

14. Une courbe dérivable $y = f(x)$ définie pour $a \leq x \leq \infty$ possède une asymptote $y = kx + b$ (chap. 4, exerc. 10) si, et seulement si, les limites

$$k = \lim_{x \rightarrow \infty} f'(x), \quad b = \lim_{x \rightarrow \infty} [f(x) - xf'(x)]$$

existent.

15. Si une fonction $f(x)$ est définie dans l'intervalle $[a, b]$ et si, dans l'intervalle (a, b) , il existe la dérivée $f'(x)$ qui a une limite p pour $x \searrow a$, alors ce nombre p est la dérivée à droite de la fonction $f(x)$ pour $x = a$.

16. Si une fonction $f(t)$ croît pour $0 \leq t \leq b$ et possède une dérivée décroissante $f'(t)$ pour $0 < t \leq b$ (la dérivée pouvant croître indéfiniment lorsque $t \rightarrow 0$), alors

$$\sum_{n=1}^{\infty} \frac{1}{n^2} f'\left(\frac{b}{n}\right) < \infty.$$

17. (Exemple de Van der Waerden.) Posons

$$\varphi_0(x) = \begin{cases} x & \text{pour } 0 \leq x \leq \frac{1}{2}, \\ 1-x & \text{pour } \frac{1}{2} \leq x \leq 1 \end{cases}$$

et prolongeons ensuite cette fonction avec 1 pour période sur tout l'axe. Posons, ensuite,

$$\varphi_n(x) = \frac{1}{4^n} \varphi_0(4^n x).$$

La fonction $\varphi_n(x)$ a la période 4^{-n} et la dérivée (partout sauf aux points anguleux d'abscisses $\frac{p}{2^n}$) qui vaut $+1$ ou -1 . Soit, enfin,

$$f(x) = \sum_{n=1}^{\infty} \varphi_n(x).$$

Montrer que $f(x)$ est continue, mais n'est dérivable en aucun point.

Historique

Le calcul différentiel (et le calcul intégral) apparaît au XVII^e siècle; considéré d'abord sous ses aspects géométrique et cinématique dans des cas particuliers par Fermat, Torricelli, Barrow, il est formulé d'une façon générale à la fin du siècle par Newton et Leibniz *). Ce dernier introduit le symbole de différentielle et la désignation $\frac{dy}{dx}$ pour la dérivée. Newton, puis Leibniz et ses disciples, les frères Jacques et Jean Bernoulli en premier lieu, appliquent les méthodes du calcul différentiel aux nombreux problèmes de géométrie, mécanique et physique. Pour les applications physiques, l'interprétation de la vitesse d'un mouvement à l'aide de la dérivée du chemin parcouru par rapport au temps joue un rôle de premier plan.

Le « théorème de Rolle » est énoncé par M. Rolle en 1690 pour les polynômes et sous forme implicite; le « théorème de Lagrange » est donné par Lagrange en 1798; les énoncés modernes de ces théorèmes n'apparaissent que dans le Cours de Cauchy (1821).

Le premier traité de calcul différentiel est écrit « en 1691 et 1692 par Johann Bernoulli à l'usage d'un marquis qui se montre bon élève » (Bourbaki). Il s'agit de Guillaume-François de L'Hospital, marquis de Sainte-Mesme, qui publie le traité mentionné en 1696, ce qui lui fait un nom dans l'histoire des sciences. Probablement, il serait plus correct du point de vue historique d'appeler règles de Bernoulli les « règles de L'Hospital ». La définition rigoureuse de la dérivée, basée sur celle de la limite, n'est donnée que par Cauchy **); depuis cette époque, « l'existence de la dérivée, au lieu d'être un article de foi, devient une question à étudier par les moyens ordinaires de l'analyse » (Bourbaki). Les exemples de fonctions continues sans dérivée sont indiqués par Bolzano en 1830 (l'ouvrage n'est publié qu'en 1930), puis par Weierstrass (1860; publication de 1872). En ce qui concerne les fondements de l'analyse, les cours de Weierstrass à l'Université de Berlin constituent une étape suivante après l'œuvre de Cauchy.

*) La première publication de Leibniz sur ces questions remonte à 1684, celle de Newton à 1687. Leur correspondance nous révèle cependant que les méthodes du calcul nouveau leur ont été familières bien avant.

**) Et, avant lui, par L'Huilier (1786); mais l'ouvrage de ce dernier n'arrive pas à propos et n'a pas de suite, bien que couronné par le prix de l'Académie des Sciences de Berlin.

Dérivées d'ordre supérieur

Si la dérivée première nous permet de décrire, à un instant donné, le caractère d'un mouvement, la dérivée seconde nous aide à pénétrer ses mobiles cachés.

W. Thomson et P. G. Tait,
Philosophia naturelle (1867)

§ 8.1. Définitions et exemples

8.11. Si une fonction numérique $y = f(x)$ admet la dérivée $y' = f'(x)$ dans un intervalle (a, b) et si la fonction $f'(x)$ est, à son tour, dérivable sur (a, b) , alors nous désignons la dérivée de la fonction $f'(x)$ par $y'' = f''(x)$ et l'appelons *dérivée seconde de la fonction $f(x)$* . En continuant, nous obtenons, sur l'intervalle (a, b) , les fonctions

$$f(x), f'(x), f''(x), \dots, f^{(n)}(x), \dots,$$

de sorte que, pour tout $n \geq 1$, $f^{(n)}(x)$ est la dérivée de $f^{(n-1)}(x)$. La fonction $f^{(n)}(x)$ s'appelle *dérivée n -ième* ou *dérivée d'ordre n de $f(x)$* ; si la fonction $f^{(n)}(x)$ existe, alors la fonction $f(x)$ est dite *n fois dérivable sur (a, b)* . Si la fonction $f'(x)$ existe et est continue, alors la fonction $f(x)$ est dite *continûment dérivable* ou *lisse*. La fonction $f(x)$ elle-même est considérée comme sa propre dérivée d'ordre 0 :

$$f(x) = f^{(0)}(x).$$

8.12. Opérations sur les dérivées d'ordre supérieur. En supposant deux fonctions $f(x)$ et $g(x)$ n fois dérivables dans un intervalle (a, b) , on a les formules suivantes :

$$(f(x) + g(x))^{(n)} = f^{(n)}(x) + g^{(n)}(x); \quad (1)$$

$$(\alpha f(x))^{(n)} = \alpha f^{(n)}(x) \quad (\alpha \text{ est un nombre réel quelconque}); \quad (2)$$

$$(f(x) \cdot g(x))^{(n)} = \sum_{h=0}^n C_n^h f^{(h)}(x) g^{(n-h)}(x) \quad (\text{formule de Leibniz}), \quad (3)$$

où $C_n^h = \frac{n!}{h!(n-h)!}$ sont des coefficients binomiaux.

On démontre les égalités (1) et (2) en dérivant successivement les égalités 7.13 (1), (2). Démontrons (3) par récurrence. Si (3) est juste pour un n , alors, en dérivant encore une fois, nous trouvons

$$\begin{aligned}(fg)^{(n+1)} &= \sum_{h=0}^n C_n^h [f^{(h)} g^{(n-h+1)} + f^{(h+1)} g^{(n-h)}] = \\&= \sum_{h=0}^n C_n^h f^{(h)} g^{(n-h+1)} + \sum_{h=1}^{n+1} C_n^{h-1} f^{(h)} g^{(n-h+1)} = \\&= fg^{(n+1)} + \sum_{h=1}^n (C_n^h + C_n^{h-1}) f^{(h)} g^{(n-h+1)} + f^{(n+1)} g = \\&= \sum_{h=0}^{n+1} C_{n+1}^h f^{(h)} g^{(n+1-h)}\end{aligned}$$

en nous basant sur la formule $C_{n+1}^h = C_n^h + C_n^{h-1}$. On voit que dès que la formule (3) est valable pour un exposant n , elle l'est pour l'exposant $n+1$; comme elle est vraie pour $n=1$, elle est établie pour n'importe quel n .

8.13. La table suivante de dérivées d'ordre supérieur pour quelques fonctions parmi les plus usuelles nous servira par la suite :

$f(x)$	$f'(x)$	$f''(x)$	$f'''(x)$	...	$f^{(n)}(x)$
x^α	$\alpha x^{\alpha-1}$	$\alpha(\alpha-1) x^{\alpha-2}$	$\alpha(\alpha-1) \times$ $\times (\alpha-2) x^{\alpha-3}$	...	$\alpha(\alpha-1) \dots \times$ $\times (\alpha-n+1) x^{\alpha-n}$
e^{ax}	$a e^{ax}$	$a^2 e^{ax}$	$a^3 e^{ax}$	...	$a^n e^{ax}$
$\ln x$	$\frac{1}{x}$	$-\frac{1}{x^2}$	$\frac{2}{x^3}$	...	$(-1)^{n-1} \frac{(n-1)!}{x^n}$
$\sin bx$	$b \cos bx$	$-b^2 \sin bx$	$-b^3 \cos bx$	...	$b^n \sin \left(bx + n \frac{\pi}{2} \right)$
$\cos bx$	$-b \sin bx$	$-b^2 \cos bx$	$+b^3 \sin bx$	...	$b^n \cos \left(bx + n \frac{\pi}{2} \right)$

8.14. Dérivées d'un polynôme et son développement suivant les puissances de $x-a$. Chaque dérivation d'un polynôme abaisse son degré d'une unité. Si un polynôme est de degré n , alors, après l'avoir n fois dérivé, nous aboutissons à une constante; donc ses dérivées d'ordre $n+1$ et plus s'annulent.

Tout polynôme $P(x)$ du n -ième degré peut être ordonné suivant les puissances croissantes de $x-a$, a étant un nombre quelconque,

par la substitution $x = (x - a) + a$; notamment, si

$$P(x) = \sum_{h=0}^n a_h x^h,$$

alors on a

$$P(x) = P[(x-a) + a] = \sum_{h=0}^n a_h [(x-a) + a]^h,$$

ce qui se met sous la forme

$$P(x) = b_0 + b_1(x-a) + b_2(x-a)^2 + \dots + b_n(x-a)^n. \quad (1)$$

Les coefficients $b_0, b_1, \dots, b_n$ s'expriment en fonction des dérivées du polynôme $P(x)$ au point a . Posant $x = a$ dans (1) on trouve

$$P(a) = b_0.$$

En dérivant (1) par rapport à x on trouve

$$P'(x) = b_1 + 2b_2(x-a) + 3b_3(x-a)^2 + \dots + nb_n(x-a)^{n-1}. \quad (2)$$

En y portant $x = a$ on obtient

$$P'(a) = b_1.$$

En dérivant (2) on obtient

$$P''(x) = 2b_2 + 2 \cdot 3b_3(x-a) + \dots + n(n-1)b_n(x-a)^{n-2}. \quad (3)$$

En y portant $x = a$ on trouve

$$P''(a) = 2b_2.$$

En continuant on obtient, pour tout $k = 0, 1, \dots, n$,

$$P^{(k)}(a) = 1 \cdot 2 \dots k \cdot b_k. \quad (4)$$

Ainsi, pour $k = 0, 1, \dots, n$, on a

$$b_k = \frac{1}{k!} P^{(k)}(a)$$

(en posant $0! = 1$), et le développement (1) se met sous la forme

$$P(x) = \sum_{h=0}^n \frac{P^{(h)}(a)}{h!} (x-a)^h. \quad (5)$$

§ 8.2. Formule de Taylor

8.21. Si une fonction $f(x)$ est continue sur un intervalle fini $[a, b]$, $n + 1$ fois dérivable dans l'intervalle (a, b) et si, de plus, ses dérivées d'ordre inférieur ou égal à n possèdent les valeurs limites $f^{(h)}(a)$ *)

*) Les limites $f^{(h)}(a)$ introduites ici sont les dérivées à droite correspondantes (exercice 15 pour le chap. 7), mais cela n'a pas d'importance.

($k = 0, 1, \dots, n$), alors il existe un point $c \in (a, b)$ tel que

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!} (b-a)^{(n+1)}. \quad (1)$$

C'est la *formule de Taylor*.

Pour $n = 0$, cette formule devient formule de Lagrange 7.44 (2). Pour un polynôme du n -ième degré, on a $f^{(n+1)}(x) \equiv 0$, et nous retrouvons la formule 8.14 (5).

L e m m e. Si deux fonctions $F(x)$ et $G(x)$ sont continues sur un intervalle $[a, b]$ et $n+1$ fois dérivables dans l'intervalle (a, b) , de sorte que $G^{(k)}(x)$ ne s'annule pas dans (a, b) et $F^{(k)}(x)$, $G^{(k)}(x)$ ont chacun 0 pour limite lorsque $x \rightarrow a$ ($k = 0, 1, \dots, n$), alors il existe un point $c \in (a, b)$ pour lequel

$$\frac{F(b)}{G(b)} = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)}.$$

En effet, d'après le théorème de Cauchy 7.43, il existe un point $c_1 \in (a, b)$ pour lequel

$$\frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(c_1)}{G'(c_1)}.$$

En appliquant le théorème de Cauchy encore une fois à l'intervalle (a, c_1) , on établit l'existence d'un point $c_2 \in (a, c_1)$ pour lequel

$$\frac{F(b)}{G(b)} = \frac{F'(c_1)}{G'(c_1)} = \frac{F'(c_1) - F'(a)}{G'(c_1) - G'(a)} = \frac{F''(c_2)}{G''(c_2)}.$$

En continuant on trouve après n pas un point $c \in (a, c_n)$ pour lequel

$$\frac{F(b)}{G(b)} = \frac{F^{(n)}(c_n)}{G^{(n)}(c_n)} = \frac{F^{(n)}(c_n) - F^{(n)}(a)}{G^{(n)}(c_n) - G^{(n)}(a)} = \frac{F^{(n+1)}(c)}{G^{(n+1)}(c)},$$

ce qu'il nous faut.

D é m o n s t r a t i o n d e l a f o r m u l e d e T a y l o r.
Posons, dans le lemme,

$$F(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k,$$

$$G(x) = (x-a)^{n+1}.$$

La fonction $F(x)$, de même que $f(x)$, possède les dérivées d'ordre inférieur ou égal à $n+1$. La fonction $G(x)$ a les dérivées de tout ordre et ses dérivées d'ordre inférieur ou égal à n sont positives pour

$x > a$. De plus, conformément à 8.14 (4),

$$\left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \right]_{x=a}^{(m)} = f^{(m)}(a),$$

de sorte que $F^{(m)}(a) = f^{(m)}(a) - f^{(m)}(a) = 0$ pour $m = 0, 1, \dots, n$. D'après la même formule $G^{(m)}(a) = 0$ pour $m = 0, 1, \dots, n$. Donc, l'hypothèse du lemme est remplie, et l'on peut l'appliquer avec

$$F^{(n+1)}(x) = f^{(n+1)}(x),$$

$$G^{(n+1)}(x) = (n+1)!$$

Le lemme appliqué, on obtient : il existe un point $c \in (a, b)$ tel que

$$\frac{F(b)}{G(b)} = \frac{f(b) - \sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!}}{(b-a)^{n+1}} = \frac{f^{(n+1)}(c)}{(n+1)!},$$

ce qui nous amène à la formule cherchée (1).

8.22. Remarque. Nous avons considéré le cas $a < b$. Dans le cas $b < a$, la démonstration se fait exactement de la même façon, de sorte que la position du point $b \neq a$ est sans importance ; dans les deux cas, le point c se trouve entre a et b .

8.23. La formule 8.21 (1) est souvent écrite comme suit :

$$f(x) = \sum_{k=0}^n f^{(k)}(a) \frac{(x-a)^k}{k!} + f^{(n+1)}(c) \frac{(x-a)^{n+1}}{(n+1)!} \quad (c \in (a, x)). \quad (1)$$

Le polynôme du n -ième degré représenté par les premiers termes du second membre s'appelle *polynôme de Taylor*. L'expression

$$R_n(x, a) = \frac{f^{(n+1)}(c)}{(n+1)!} (x-a)^{n+1} \quad (2)$$

est le *reste de la formule de Taylor* (sous forme de Lagrange). Si $f^{(n+1)}(x)$ est borné pour $x \rightarrow a$, alors le reste R_n est un infiniment petit par rapport à $(x-a)^n$ lorsque $x \rightarrow a$ (4.38) ; conformément à la désignation adoptée dans 4.38, nous notons

$$R_n = o((x-a)^n).$$

En particulier, il est aisé de prouver que, pour $x < 0$ ou $x > 0$, les développements suivants ont lieu :

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + o(x^n); \quad (3)$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}); \quad (4)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}); \quad (5)$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{n-1} \frac{x^n}{n} + o(x^n); \quad (6)$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 + \dots + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{1 \cdot 2 \dots n} x^n + o(x^n). \quad (7)$$

8.24. La formule de Taylor sert à représenter une fonction $f(x)$ (qui peut être assez compliquée) par une fonction relativement simple, notamment par un polynôme, à 8.23 (2) près, l'erreur permettant, dans certains cas, une estimation simple et pouvant être rendue assez petite. Si l'on sait que

$$\sup_{a \leq x \leq a+h} |f^{(n)}(x)| \leq M_n(h) \quad (n = 0, 1, 2, \dots),$$

alors on obtient l'estimation suivante pour le reste 8.23 (2) :

$$|R_n| \leq \frac{M_{n+1}(h)}{(n+1)!} h^{n+1}. \quad (1)$$

Désignons le second membre par $\omega = \omega(n, h)$. On peut formuler trois problèmes naturels liés à la détermination de chacune des quantités ω , n , h en fonction de deux autres :

(a) Pour n et h donnés, trouver $\omega(n, h)$; en d'autres termes, trouver un majorant pour l'erreur commise en remplaçant la fonction $f(x)$ dans l'intervalle $(a, a+h)$ par son polynôme de Taylor du n -ième degré.

(b) Pour n et ω donnés, trouver h ; en d'autres termes, trouver un intervalle $(a, a+h)$ sur lequel l'erreur commise en remplaçant la fonction $f(x)$ par son polynôme de Taylor du n -ième degré ne dépasse pas la valeur donnée ω .

(c) Pour h et ω donnés, trouver n ; en d'autres termes, trouver le degré d'un polynôme de Taylor, si l'erreur commise en remplaçant par ce polynôme la fonction $f(x)$ dans un intervalle donné $(a, a+h)$ ne dépasse pas la valeur donnée ω .

Dans des cas concrets, tous ces problèmes se résolvent d'une façon plus ou moins élémentaire (voir l'exercice 10).

On peut formuler les mêmes problèmes pour un intervalle $(a-h, a)$, les méthodes de résolution restant les mêmes.

§ 8.3. Analyse du comportement d'une fonction au voisinage d'un point donné

8.31. Dans § 7.5, nous avons discuté la position d'une courbe $y = f(x)$ par rapport à sa tangente au $x = x_0$. L'appareil analytique était en l'occurrence les valeurs de $f'(x)$ pour $x \neq x_0$. Revenons à ce problème en utilisant, au lieu d'un ensemble de valeurs de $f'(x)$, une seule valeur de $f''(x)$ pour $x = x_0$.

T h é o r è m e. Soit $y = f(x)$ une fonction qui possède la dérivée première $f'(x)$ dans un intervalle (a, b) contenant un point x_0 et la dérivée seconde $f''(x_0)$ au point x_0 . Si $f''(x_0) > 0$, alors x_0 est un point de convexité dirigée vers le bas pour la fonction $f(x)$; si $f''(x_0) < 0$, alors x_0 est un point de convexité dirigée vers le haut pour cette fonction.

D é m o n s t r a t i o n. Soit $f''(x_0) > 0$; en appliquant le résultat 7.23.a à la fonction $f'(x)$, on voit qu'il existe un $\delta > 0$ tel que, pour $0 < h < \delta$, on a :

$$f'(x_0 - h) < f'(x_0) < f'(x_0 + h).$$

Ceci permet d'appliquer le résultat 7.54 à $(x_0 - \delta, x_0 + \delta)$, ce qui achève la démonstration. Le cas $f''(x_0) < 0$ est analysé de façon analogue.

8.32. Le résultat obtenu a pour conséquences les nouvelles conditions suffisantes d'existence d'un minimum et d'un maximum.

a. Si, dans l'hypothèse de 8.31, nous avons $f'(x_0) = 0$ et $f''(x_0) > 0$, alors le point x_0 est un point de minimum local pour la fonction $f(x)$.

b. Si, dans l'hypothèse de 8.31, nous avons $f'(x_0) = 0$ et $f''(x_0) < 0$, alors le point x_0 est un point de maximum local pour la fonction $f(x)$.

Les deux propositions résultent de 8.31 et 7.56.

8.33. Si l'on connaît les valeurs de $f''(x)$ non seulement au point x_0 mais aussi dans un voisinage de ce point, quel supplément d'information sur les propriétés de la fonction $f(x)$ peut-on en tirer? Il s'avère que l'on est alors en mesure de décrire la position de la courbe $y = f(x)$ au voisinage du point x_0 par rapport à sa « parabole osculatrice »

$$Y(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2. \quad (1)$$

a. **T h é o r è m e.** Lorsque, pour tous les $\eta \in (x_0, b)$, on a l'inégalité $f''(\eta) > f''(x_0)$, alors la courbe $y = f(x)$, pour $x > x_0$, est au-dessus de la parabole osculatrice (1). Lorsque, pour tous les $\eta \in$

$\in (x_0, b)$, on a l'inégalité $f''(\eta) < f''(x_0)$, alors la courbe $y = f(x)$, pour $x > x_0$, est au-dessous de la parabole (1).

Démonstration. Dans la première hypothèse, on a pour $x > x_0$, d'après la formule de Taylor avec $n = 2$:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(c)(x - x_0)^2 > \\ &> f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 = Y(x). \end{aligned}$$

Dans la seconde hypothèse, on raisonne de la même manière.

b. D'une façon analogue, lorsque, pour tous les $\xi \in (a, x_0)$, on a l'inégalité $f''(\xi) > f''(x_0)$, alors la courbe $y = f(x)$, pour $x < x_0$, est au-dessus de la parabole (1); lorsque, pour tous les $\xi \in (a, x_0)$, on a l'inégalité $f''(\xi) < f''(x_0)$, alors la courbe $y = f(x)$, pour $x < x_0$, est au-dessous de la parabole (1).

8.34. Envisageons, dans les conditions 8.33, le cas où $f''(x_0) = 0$. Alors la parabole osculatrice 8.33 (1) dégénère en la droite tangente à la courbe $y = f(x)$ pour $x = x_0$:

$$Y(x) = f(x_0) + f'(x_0)(x - x_0).$$

En combinant les résultats 8.33 on aboutit aux conclusions suivantes:

Théorème. Si, dans les conditions 8.33, on a $f''(x_0) = 0$ et que, pour tous les $x \neq x_0$, on a l'inégalité $f''(x) > f''(x_0)$, alors le point x_0 est un point de convexité dirigée vers le bas pour la fonction $f(x)$. Si l'on a l'inégalité $f''(x) < f''(x_0)$ pour tous les $x \neq x_0$, alors le point x_0 est un point de convexité dirigée vers le haut pour la fonction $f(x)$. Si, pour tous les $\xi \in (a, x_0)$, $\eta \in (x_0, b)$, ont lieu les inégalités

$$f''(\xi) < 0 < f''(\eta),$$

ou bien les inégalités

$$f''(\xi) > 0 > f''(\eta),$$

alors le point x_0 est un point d'inflexion pour la fonction $f(x)$.

Soulignons que les deux dernières suppositions sont des conditions suffisantes pour que le point x_0 soit un point d'inflexion.

8.35. Tout comme pour $f'(x)$ plus haut, on peut utiliser, au lieu d'un ensemble de valeurs de $f''(x)$, une seule valeur de $f'''(x)$ au point x_0 . Supposons que la fonction $f(x)$ possède les dérivées $f'(x)$ et $f''(x)$ dans un intervalle $(a, b) \ni x_0$ et la dérivée $f'''(x_0)$ au point x_0 .

Théorème. Si $f'''(x_0) > 0$, alors la courbe $y = f(x)$ est au-dessus de la parabole 8.33 (1) pour $x > x_0$ et au-dessous pour $x <$

$< x_0$. Si $f'''(x_0) < 0$, la courbe est par contre au-dessus de la parabole 8.33 (1) pour $x < x_0$ et au-dessous pour $x > x_0$.

Démonstration. Soit $f'''(x_0) > 0$. En appliquant le résultat 7.23.a à la fonction $f''(x)$ on voit qu'il existe un $\delta > 0$ tel que l'inégalité

$$f''(x_0 - h) < f''(x_0) < f''(x_0 + h)$$

a lieu pour $0 < h < \delta$. On peut maintenant appliquer les théorèmes 8.33.a, b à l'intervalle $(x_0 - \delta, x_0 + \delta)$, ce qui nous conduit aux résultats cherchés.

8.36. Dans le cas où $f''(x_0) = 0$, on obtient: si $f''(x_0) = 0$, $f'''(x_0) \neq 0$, alors x_0 est un point d'inflexion de la fonction $f(x)$.

Dans le cas où $f'''(x_0) = 0$, la situation reste indéterminée (une fois que l'on ne veut utiliser les valeurs de la fonction et de ses dérivées qu'au seul point x_0). On peut se servir des valeurs de dérivées suivantes (voir l'exercice 11).

8.37. Une autre forme de la formule de Taylor qui utilise le symbole E (4.39) est utile, lorsqu'on étudie le comportement d'une fonction au voisinage d'un point donné. Supposons qu'une fonction $f(x)$ possède les dérivées successives, y compris la $n+1$ -ième, au voisinage d'un point a , de sorte que $f^{(n)}(a) \neq 0$ et que $f^{(n+1)}(x)$ soit bornée. Alors la formule de Taylor 8.23 (1) peut être mise sous la forme:

$$f(a+h) = f(a) + hf'(a) + \dots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(a) + \frac{h^n}{n!} f^{(n)}(a) E, \quad (1)$$

où, comme plus haut, E désigne une quantité qui tend vers 1 (pour $h \rightarrow 0$). Dans le présent cas on a, évidemment:

$$E = 1 + \frac{h}{n+1} \frac{f^{(n+1)}(c)}{f^{(n)}(a)}.$$

Signalons en particulier les formules qui résultent de (1) (pour les valeurs de n données ci-dessous) et de 8.13 (ou bien 8.23 (3)-(7)):

$$e^x = 1 + xE \quad (n=1); \quad (2)$$

$$\cos x = 1 - \frac{x^2}{2} E \quad (n=2); \quad (3)$$

$$\sin x = xE \quad (n=1); \quad (4)$$

$$\ln(1+x) = xE \quad (n=1); \quad (5)$$

$$(1+x)^\alpha = 1 + \alpha xE \quad (n=1); \quad (6)$$

quant à la dernière formule, nous l'avons obtenue dans 5.59.f.

Des formules plus exactes qui résultent toujours de (1) et de 8.23 (3)-(7) sont :

$$e^x = 1 + x + \frac{x^2}{2} E \quad (n=2); \quad (2')$$

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} E \quad (n=4); \quad (3')$$

$$\sin x = x - \frac{x^3}{6} E \quad (n=3); \quad (4')$$

$$\ln(1+x) = x + \frac{x^2}{2} E \quad (n=2); \quad (5')$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 E \quad (n=2). \quad (6')$$

8.38. Exemples.

a. Trouver

$$\lim_{x \rightarrow 0} \frac{e^x - \cos x}{\sin x}.$$

Solution. En appliquant les formules 8.37 (2)-(4) on trouve :

$$\frac{e^x - \cos x}{\sin x} = \frac{(1+xE) - \left(1 - \frac{x^2}{2} E\right)}{xE} = \frac{E + \frac{x}{2} E}{E} \rightarrow 1.$$

(On aurait pu aussi bien appliquer la règle de L'Hospital 7.61.)

b. Trouver

$$\lim_{x \rightarrow 0} \frac{e^{x^2} - \sqrt{1-x^2+x^3}}{\ln(1+x^2)}.$$

Solution. En se servant des formules 8.37 (2), (5), (6) on trouve

$$\begin{aligned} \frac{e^{x^2} - \sqrt{1-x^2+x^3}}{\ln(1+x^2)} &= \frac{1+x^2E - \left(1 - \frac{x^2-x^3}{2} E\right)}{x^2E} = \\ &= \frac{\frac{3}{2}E - \frac{x}{2}E}{E} \rightarrow \frac{3}{2}. \end{aligned}$$

Ici l'application de la règle de L'Hospital conduirait à des calculs compliqués.

c. Quelle est la position de la courbe

$$y = f(x) = \frac{x - \sin x}{x^3}$$

au voisinage du point $x=0$?

Solution. D'après la formule 8.37 (4'),

$$y = \frac{x - \left(x - \frac{x^3}{6} E\right)}{x^3} = \frac{1}{6} E.$$

de sorte que $\lim_{x \rightarrow 0} y = \frac{1}{6}$. Une réponse plus exacte exige l'application d'une formule plus exacte pour $\sin x$, notamment

$$\sin x = x - \frac{x^3}{6} + \frac{x^5}{120} E$$

(ce qui résulte de 8.23 (5)). On a alors

$$y = \frac{1}{6} - \frac{x^2}{120} E,$$

de sorte que la fonction $y = f(x)$ a un maximum local au point $x = 0$.

§ 8.4. Différentielles d'ordre supérieur

8.41. Supposons qu'une fonction $y = f(x)$ soit définie dans un intervalle (a, b) et y possède les dérivées successives, y compris la n -ième. On appelle *différentielle d'ordre n de la fonction $f(x)$* une fonction de deux variables indépendantes x et dx qui est définie par l'égalité

$$d^n y \equiv d^n f(x) = f^{(n)}(x) (dx)^n.$$

Lorsque dx est constant, alors, évidemment, $d^n y = d(d^{n-1}y)$. Etant donné $d^n y$, la dérivée d'ordre n peut être mise sous la forme d'un rapport :

$$f^{(n)}(x) = \frac{d^n y}{(dx)^n}.$$

En particulier, pour $y = f(x) \equiv x$, nous avons

$$dy = dx, d^2 y = d^3 y = \dots = 0.$$

8.42. Explicitons la différentielle seconde d'une fonction composée $y = f[\varphi(x)]$. D'après la définition 8.41, on a

$$\begin{aligned} d^2 y &= \{f[\varphi(x)]\}'' (dx)^2 = [f'[\varphi] \cdot \varphi'(x)]' (dx)^2 = \\ &= f''[\varphi(x)] \cdot [\varphi'(x) dx]^2 + f'[\varphi] \cdot \varphi''(x) (dx)^2 = \\ &= f''[\varphi(x)] \cdot (d\varphi)^2 + f'[\varphi] \varphi''(x) (dx)^2. \end{aligned} \quad (1)$$

Si φ était une variable indépendante, on aurait

$$d^2 y = f''[\varphi] (d\varphi)^2.$$

Nous voyons que, contrairement à la différentielle première, la différentielle seconde (et suivantes) changent de forme lorsque la variable indépendante devient dépendante, i.e. fonction d'une nouvelle variable indépendante. D'ailleurs, dans le cas où la variable indépendante ne subit qu'une transformation linéaire, la différentielle seconde et les suivantes ne changent pas de forme. Dans ce cas,

$\varphi(x) = Ax + B$ (A et B sont constants), donc $\varphi''(x) = 0$ et le deuxième terme dans (1) disparaît. En effectuant les différentiations suivantes, on obtient successivement :

$$d^3y = f''' [\varphi(x)] [\varphi'(x) dx]^3 = f''' [\varphi(x)] (d\varphi)^3,$$

$$\dots \dots \dots$$

$$d^ny = f^{(n)} [\varphi(x)] [\varphi'(x) dx]^n = f^{(n)} [\varphi(x)] (d\varphi)^n,$$

car les termes contenant les dérivées supérieures de φ s'annulent.

8.43. Avec les différentielles, la formule de Taylor 8.21 (1) prend la forme suivante ($u = x$, $b - u = dx$) :

$$f(x + dx) = \sum_{k=0}^n \frac{1}{k!} d^k f(u) + \frac{1}{(n+1)!} d^{n+1} f(c).$$

§ 8.5. Série de Taylor

8.51. Considérons une fonction $y = f(x)$ qui possède, dans un intervalle (a, b) , les dérivées de tous les ordres $n = 0, 1, 2, \dots$. Dans la suite, de telles fonctions seront dites *indéfiniment dérivables*. Choisissons un point $x_0 \in (a, b)$ et écrivons la formule de Taylor :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n(x, x_0).$$

Supposons que, pour un x donné, la valeur du reste tende vers 0 lorsque n croît. Cela veut dire que la série infinie

$$\begin{aligned} f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + \dots = \\ = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \end{aligned} \quad (1)$$

converge et a pour somme le nombre $f(x)$. Donc, la valeur de la fonction $f(x)$ pour un x donné est la somme de la série de puissances de $x - x_0$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + \dots$$

Indépendamment de sa convergence et de sa somme, la série (1) pour une fonction indéfiniment dérivable $f(x)$ s'appelle *série de Taylor relative à la fonction $f(x)$* . Elucidons pour quelles fonctions $f(x)$ et dans quels domaines la série de Taylor converge et a pour somme la valeur de la fonction $f(x)$ elle-même ; en d'autres termes, pour quelles fonctions et dans quels domaines le reste de la formule de Taylor tend vers 0.

8.52. *L e m m e. Si, dans un intervalle (a, b) et pour certains $B > 0$ et $C > 0$, on a l'inégalité*

$$\frac{1}{B^n \cdot n!} \sup_{x \in (a, b)} |f^{(n)}(x)| \leq C \quad (n = 0, 1, 2 \dots) \quad (1)$$

(i.e. que la suite numérique du premier membre est bornée), alors le reste de la formule de Taylor tend vers 0 en tout point $x \in (a, b)$

pour lequel $|x - x_0| < \frac{1}{B}$.

Démonstration. Expliciteons le reste :

$$R_n(x, x_0) = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(c) \quad (c \text{ entre } x \text{ et } x_0).$$

En vertu de (1), nous avons

$$|R_n(x, x_0)| \leq \frac{|x - x_0|^{n+1}}{(n+1)!} C B^{n+1} (n+1)! = C (B |x - x_0|)^{n+1},$$

qui tend vers 0, lorsque $n \rightarrow \infty$, une fois que $|x - x_0| < \frac{1}{B}$.

8.53. Considérons quelques exemples.

a. Pour un polynôme $y = f(x)$ de degré m , l'inégalité 8.52 (1) est immédiate, car $f^{(m+1)}(x)$ et toutes les dérivées suivantes sont nulles. Donc tout polynôme se développe en série de Taylor; d'ailleurs, cette série se réduit à une somme finie que nous avons déjà vue dans 8.14.

b. Soit $f(x) = e^x$. Nous avons $f^{(n)}(x) = e^x$ et

$$\sup_{|x| \leq b} e^x = e^b$$

dans l'intervalle $[-b, b]$. Par conséquent, l'expression

$$\frac{\sup |f^{(n)}(x)|}{n! B^n} \leq \frac{e^b}{n! B^n}$$

tend vers 0 lorsque n croît, donc reste bornée pour $n \rightarrow \infty$ quel que soit un b fixe. Ainsi, la série de Taylor de la fonction e^x converge vers e^x sur l'intervalle $[-b, b]$ de n'importe quelle longueur; autrement dit, elle converge vers e^x pour tous les x réels. Donc, pour tous les x , on a

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots \quad (1)$$

c. Soit $f(x) = \sin x$. Nous avons $|f^{(n)}(x)| \leq 1$ et

$$\frac{\sup |f^{(n)}(x)|}{B^n \cdot n!} \leq \frac{1}{B^n \cdot n!}$$

dans tout intervalle $[-b, b]$. Cette suite est bornée indépendamment de b . Donc, la fonction $\sin x$ se développe elle aussi en série de Taylor qui converge sur tout l'axe réel :

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (2)$$

d. De même, la fonction $\cos x$ se développe en série de Taylor qui converge sur tout l'axe réel :

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (3)$$

8.54. Notons que l'existence des développements 8.53 (2) et (3) démontre l'unicité des fonctions $\sin x$ et $\cos x$ que nous avons définies dans 5.61 comme solutions de certaines équations fonctionnelles. Pour achever leur théorie, il nous faut, en partant des développements 8.53 (2), (3), montrer que les équations et les inégalités mentionnées dans 5.61 sont satisfaites; on le fera dans 8.68.

8.55. **Prolongement analytique.** Supposons qu'une fonction $f(x)$, indéfiniment dérivable dans un intervalle (a, b) , s'y développe en série de Taylor

$$f(x) = \sum_{k=0}^{\infty} c_k (x-x_0)^k, \quad c_k = \frac{f^{(k)}(x_0)}{k!}. \quad (1)$$

Considérons cette série dans le domaine complexe en remplaçant la variable réelle x par la variable complexe $z = x + iy$:

$$\sum_{k=0}^{\infty} c_k (z-x_0)^k. \quad (2)$$

Le domaine de convergence de cette série est le cercle d'un rayon R centré au point x_0 (6.52). De toute façon, le nombre R est tel que tout l'intervalle (a, b) dans lequel la série (1) converge est inclus dans le cercle $|z - x_0| < R$, puisque l'on sait qu'une série de puissances diverge en tout point extérieur à son cercle de convergence. En tout point $z \in (a, b)$, la somme de la série (2) vaut $f(x)$ par hypothèse. La somme de la série (2) considérée en d'autres points du cercle de convergence représente une fonction de z ; nous l'appellerons *prolongement analytique* de la fonction $f(x)$ dans le domaine complexe et désignerons par $f(z)$. En particulier, si la série (1) converge pour tous les x réels, alors la série (2) converge pour tous les z complexes, le prolongement analytique de la fonction $f(x)$ se trouvant ainsi défini dans tout le plan complexe.

§ 8.6. Exponentielle et fonctions trigonométriques dans le domaine complexe

Appliquons les résultats 8.55 d'ordre général à l'exponentielle et aux fonctions trigonométriques.

8.61. La série de Taylor de la fonction e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

converge pour tous les x réels; donc la série

$$e^z = 1 + z + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots \quad (1)$$

converge pour tous les z complexes en définissant l'exponentielle e^z dans tout le plan de la variable z .

8.62. D'une façon analogue, la série de Taylor de la fonction $\sin x$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots,$$

ainsi que celle de la fonction $\cos x$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots,$$

converge pour tous les x réels; par conséquent, pour tous les z complexes, sont définies les fonctions

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots, \quad (1)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots \quad (2)$$

qui sont les prolongements analytiques des fonctions trigonométriques $\sin x$ et $\cos x$.

On voit des formules (1) et (2) que, dans le plan complexe (comme sur l'axe réel), on a les égalités

$$\sin(-z) = -\sin z, \quad (3)$$

$$\cos(-z) = \cos z, \quad (4)$$

qui expriment respectivement l'imparité de la fonction $\sin z$ et la parité de $\cos z$ dans le plan complexe.

8.63. Une relation remarquable qui existe entre l'exponentielle et les fonctions trigonométriques dans le domaine complexe résulte des développements 8.61 (1) et 8.62 (1), (2). Notamment, en rempla-

çant dans 8.61 (1) z par iz , on trouve que, pour tout z ,

$$e^{iz} = 1 + iz - \frac{z^2}{2!} - i \frac{z^3}{3!} + \frac{z^4}{4!} + \dots = \\ = \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots\right) + i \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots\right) = \cos z + i \sin z. \quad (1)$$

Ainsi, l'exponentielle est une combinaison linéaire de $\sin z$ et $\cos z$. En remplaçant dans (1) z par $-z$ on obtient encore (compte tenu de 8.62 (3) et (4))

$$e^{-iz} = \cos z - i \sin z. \quad (2)$$

Il résulte de (1) et (2):

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad (3)$$

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}. \quad (4)$$

D'ailleurs, ces formules auraient pu être obtenues directement de 8.61 (1) et 8.62 (1), (2). Les formules (1)-(4) s'appellent *formules d'Euler*.

Pour un z réel qui vaut θ , nous avons de (1):

$$e^{i\theta} = \cos \theta + i \sin \theta. \quad (5)$$

On voit que le module du nombre complexe $e^{i\theta}$, avec θ réel, est égal à l'unité.

Pour tout nombre complexe $z \neq 0$, nous avons eu, dans 5.72, la représentation trigonométrique

$$z = r(\cos \theta + i \sin \theta),$$

où θ est l'argument et r le module de z . L'égalité (5) permet de mettre le même nombre z sous « forme exponentielle »

$$z = re^{i\theta}. \quad (6)$$

8.64. Les autres propriétés de l'exponentielle sont basées sur le fait suivant:

L e m m e. *Quels que soient z_1 et z_2 complexes, on a l'égalité*

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}.$$

En effet, on a

$$e^{z_1} = 1 + z_1 + \frac{z_1^2}{2!} + \dots,$$

$$e^{z_2} = 1 + z_2 + \frac{z_2^2}{2!} + \dots$$

Les deux séries étant absolument convergentes, on peut appliquer le théorème de multiplication (6.46) qui donne

$$\begin{aligned} e^{z_1} \cdot e^{z_2} &= 1 + (z_1 + z_2) + \left(\frac{z_1^2}{2!} + z_1 z_2 + \frac{z_2^2}{2!} \right) + \\ &+ \left(\frac{z_1^3}{3!} + \frac{z_1^2 z_2}{2! \cdot 1} + \frac{z_1 z_2^2}{1 \cdot 2!} + \frac{z_2^3}{3!} \right) + \dots = \\ &= 1 + (z_1 + z_2) + \frac{(z_1 + z_2)^2}{2!} + \frac{(z_1 + z_2)^3}{3!} + \dots = e^{z_1 + z_2}, \end{aligned}$$

ce qu'il fallait démontrer.

8.65. En particulier, pour x et y réels et pour $z = x + iy$, on a

$$e^z = e^{x+iy} = e^x \cdot e^{iy} = e^x (\cos y + i \sin y) = e^x \cos y + i e^x \sin y, \quad (1)$$

formule qui explicite les parties réelle et imaginaire de la fonction e^z .

La relation (1) révèle une particularité intéressante de la fonction e^z dans le plan complexe, à savoir sa périodicité. D'une façon générale, une fonction $f(z)$ s'appelle *fonction périodique de période (complexe) T* si la relation

$$f(z + T) = f(z)$$

est satisfaite pour tout z du domaine de définition de la fonction $f(z)$.

En particulier, pour $f(z) = e^z$ et $T = 2\pi i$, nous avons d'après (1) :

$$e^{z+2\pi i} = e^{x+iy+2\pi i} = e^{x+i(y+2\pi)} =$$

$$= e^x [\cos(y+2\pi) + i \sin(y+2\pi)] = e^x [\cos y + i \sin y] = e^z, \quad (2)$$

de sorte que la fonction $f(z) = e^z$ s'avère *périodique de période purement imaginaire $T = 2\pi i$* . Donc, pour tout $k = 0, \pm 1, \pm 2, \dots$:

$$e^{z+2\pi k i} = e^z.$$

Au contraire, les fonctions e^{iz} et e^{-iz} ont pour période, dans le plan complexe, le nombre réel $T = 2\pi$. Par exemple, pour $f(z) = e^{iz}$, nous tirons de (2) en y remplaçant z par iz :

$$f(z+2\pi) = e^{i(z+2\pi)} = e^{iz+2\pi i} = e^{iz} = f(z).$$

En vertu des formules 8.63 (3), (4), les fonctions $\sin z$ et $\cos z$ considérées dans tout le plan complexe sont périodiques de période réelle 2π , de même que e^{iz} et e^{-iz} .

En posant $z = i\pi$ dans (1) on aboutit à une formule bien intéressante *) :

$$e^{i\pi} = -1.$$

*) « On dirait que cette formule remarquable d'Euler symbolise l'unité des mathématiques : -1 y représente l'arithmétique, i l'algèbre, π la géométrie et e l'analyse » (A. Krylov).

8.66. D'autre part, en multipliant membre à membre les relations 8.63 (1) et (2) et en appliquant le lemme 8.64 on trouve

$$e^{iz} \cdot e^{-iz} = e^0 = 1 = (\cos z + i \sin z) (\cos z - i \sin z) = \cos^2 z + \sin^2 z. \quad (1)$$

Ainsi, les fonctions $\cos z$ et $\sin z$ satisfont, dans le plan complexe aussi, à la première des équations fonctionnelles qui nous ont servi à définir ces fonctions (5.61 (1)). Cependant, la conclusion que l'on a tirée de l'équation (1) dans le domaine réel et qui dit que les fonctions $\sin z$ et $\cos z$ sont bornées en module par le nombre 1, n'en résulte plus, car les valeurs de ces fonctions sont elles-mêmes complexes. Qui plus est, les relations 8.63 (3), (4) où l'on remplace z par iy montrent que les fonctions $\cos(iy)$ et $i \sin(iy)$ (leur argument est purement imaginaire si y est un nombre réel) croissent indéfiniment pour $y \rightarrow \pm\infty$.

8.67. Soient z_1 et z_2 des nombres complexes quelconques. En remplaçant dans les développements 8.63 (1) et (2) z par z_1 et par z_2 et en multipliant on trouve

$$\begin{aligned} e^{i(z_1+z_2)} &= \cos(z_1 + z_2) + i \sin(z_1 + z_2) = e^{iz_1} \cdot e^{iz_2} = \\ &= (\cos z_1 + i \sin z_1) (\cos z_2 + i \sin z_2) = \\ &= (\cos z_1 \cos z_2 - \sin z_1 \sin z_2) + i (\cos z_1 \sin z_2 + \sin z_1 \cos z_2) \end{aligned}$$

et, d'une manière analogue,

$$\begin{aligned} e^{-i(z_1+z_2)} &= \cos(z_1 + z_2) - i \sin(z_1 + z_2) = e^{-iz_1} \cdot e^{-iz_2} = \\ &= (\cos z_1 - i \sin z_1) (\cos z_2 - i \sin z_2) = \\ &= (\cos z_1 \cos z_2 - \sin z_1 \sin z_2) - i (\cos z_1 \sin z_2 + \sin z_1 \cos z_2). \end{aligned}$$

En additionnant et soustrayant les développements obtenus on trouve

$$\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2, \quad (1)$$

$$\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2, \quad (2)$$

ce qui prolonge les équations fonctionnelles 5.61 (2) et (3) dans le plan complexe.

8.68. En particulier, si l'on prend les séries 8.62 (1) et (2) pour définitions des fonctions $\sin x$ et $\cos x$, on voit que les équations fonctionnelles 5.61 (2) et (3) sont satisfaites. Montrons que, pour x positif et assez petit, il en est de même de l'inégalité 5.61 (4)

$$\sin x < x < \frac{\sin x}{\cos x}. \quad (1)$$

En partant de 8.62 (1) on conclut que

$$\sin x = x - \frac{x^3}{6} E = x \left(1 - \frac{x^2}{6} E \right), \quad (2)$$

où E tend vers 1 pour $x \rightarrow 0$. Donc, pour x assez petit, on a $E > \frac{1}{2}$, et l'expression entre parenthèses dans (2) est inférieure à 1, d'où la première inégalité (1).

Ensuite, le développement 8.62 (2) implique la représentation

$$\cos x = 1 - \frac{x^2}{2} E, \quad (3)$$

où $E \rightarrow 1$ lorsque $x \rightarrow 0$. Par conséquent (vu 8.37 (6)),

$$\begin{aligned} \frac{\sin x}{\cos x} &= x \left(1 - \frac{x^2}{6} E\right) \left(1 - \frac{x^2}{2} E\right)^{-1} = \\ &= x \left(1 - \frac{x^2}{6} E\right) \left(1 + \frac{x^2}{2} E\right) = x \left(1 + \frac{x^2}{3} E\right). \end{aligned} \quad (4)$$

Ici l'on a $E > \frac{1}{2}$ pour x assez petit, et l'expression entre parenthèses devient supérieure à 1, ce qui démontre notre proposition.

De la sorte, les fonctions vérifiant les relations 5.61 (1)-(4) peuvent être définies par les séries 8.62 (1) et (2). L'unicité des solutions de ces équations a déjà été établie (8.54).

§ 8.7. Fonctions hyperboliques

8.71. Définissons deux fonctions d'une variable complexe :

$$\operatorname{ch} z = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \frac{z^6}{6!} + \dots, \quad (1)$$

$$\operatorname{sh} z = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots \quad (2)$$

Les séries convergent pour tout z complexe. La première des fonctions définies s'appelle *cosinus hyperbolique de z* , la deuxième *sinus hyperbolique de z* . Elles sont désignées respectivement par $\operatorname{ch} z$ et $\operatorname{sh} z$.

8.72. En partant du développement 8.61 (1)

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$$

et de la relation que l'on obtient en y remplaçant z par $-z$:

$$e^{-z} = 1 - z + \frac{z^2}{2!} - \frac{z^3}{3!} + \dots,$$

on peut exprimer les exponentielles e^z et e^{-z} par les fonctions introduites $\operatorname{ch} z$ et $\operatorname{sh} z$:

$$\begin{aligned} e^z &= \operatorname{ch} z + \operatorname{sh} z, \\ e^{-z} &= \operatorname{ch} z - \operatorname{sh} z. \end{aligned}$$

En résolvant les équations obtenues par rapport à $\operatorname{ch} z$ et $\operatorname{sh} z$ on trouve

$$\operatorname{ch} z = \frac{1}{2} (e^z + e^{-z}), \quad (1)$$

$$\operatorname{sh} z = \frac{1}{2} (e^z - e^{-z}). \quad (2)$$

En particulier, pour un x réel, on a

$$\operatorname{ch} x = \frac{1}{2} (e^x + e^{-x}), \quad (3)$$

$$\operatorname{sh} x = \frac{1}{2} (e^x - e^{-x}). \quad (4)$$

Les graphiques de ces fonctions sont donnés fig. 8.1.

8.73. Dans le domaine complexe, les fonctions hyperboliques sont liées d'une façon étroite aux fonctions trigonométriques. Notamment, en remplaçant dans 8.72 (1) et (2) z par iz , on trouve

$$\cos z = \frac{1}{2} (e^{iz} + e^{-iz}) = \operatorname{ch} iz, \quad (1)$$

$$i \sin z = \frac{1}{2} (e^{iz} - e^{-iz}) = \operatorname{sh} iz. \quad (2)$$

En remplaçant encore une fois iz par z , on trouve

$$\operatorname{ch} z = \cos(-iz) = \cos iz, \quad (3)$$

$$\operatorname{sh} z = i \sin(-iz) = -i \sin iz. \quad (4)$$

Ainsi, les fonctions trigonométriques se déduisent des fonctions hyperboliques en remplaçant la variable z par iz (au facteur i devant $\sin z$ près). Géométriquement, la multiplication de la variable par i correspond à la rotation de 90° dans le plan (5.72).

Signalons spécialement les formules (x et y sont réels)

$$\begin{aligned} \sin(x + iy) &= \sin x \cdot \cos iy + \cos x \cdot \sin iy = \\ &= \sin x \cdot \operatorname{ch} y + i \cos x \cdot \operatorname{sh} y, \end{aligned} \quad (5)$$

$$\begin{aligned} \cos(x + iy) &= \cos x \cdot \cos iy - \sin x \cdot \sin iy = \\ &= \cos x \cdot \operatorname{ch} y - i \sin x \cdot \operatorname{sh} y. \end{aligned} \quad (6)$$

Ces formules se déduisent de (1) et (2) et de 8.67 (1) et (2).

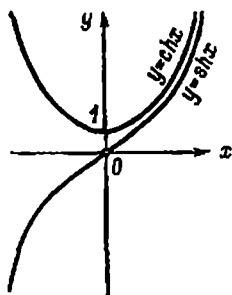


Fig. 8.1.

8.74. Nous avons encore :

$$\operatorname{ch}^2 z - \operatorname{sh}^2 z = \cos^2 (iz) + \sin^2 (iz) = 1; \quad (1)$$

$$\begin{aligned} \operatorname{ch} (z_1 + z_2) &= \cos i (z_1 + z_2) = \cos iz_1 \cos iz_2 - \\ &\quad - \sin iz_1 \sin iz_2 = \operatorname{ch} z_1 \operatorname{ch} z_2 + \operatorname{sh} z_1 \operatorname{sh} z_2; \end{aligned} \quad (2)$$

$$\begin{aligned} \operatorname{sh} (z_1 + z_2) &= -i \sin i (z_1 + z_2) = \\ &= -i [\sin iz_1 \cos iz_2 + \cos iz_1 \sin iz_2] = \\ &= \operatorname{sh} z_1 \operatorname{ch} z_2 + \operatorname{ch} z_1 \operatorname{sh} z_2. \end{aligned} \quad (3)$$

En particulier,

$$\operatorname{ch} 2z = \operatorname{ch}^2 z + \operatorname{sh}^2 z, \quad (4)$$

$$\operatorname{sh} 2z = 2 \operatorname{sh} z \operatorname{ch} z. \quad (5)$$

Ensuite, pour x et y réels,

$$\begin{aligned} \operatorname{sh} (x + iy) &= \operatorname{sh} x \operatorname{ch} iy + \operatorname{ch} x \operatorname{sh} iy = \\ &= \operatorname{sh} x \cos y + i \operatorname{ch} x \sin y, \end{aligned} \quad (6)$$

$$\begin{aligned} \operatorname{ch} (x + iy) &= \operatorname{ch} x \operatorname{ch} iy + \operatorname{sh} x \operatorname{sh} iy = \\ &= \operatorname{ch} x \cos y + i \operatorname{sh} x \sin y. \end{aligned} \quad (7)$$

8.75. Signalons, enfin, les formules de dérivation pour x réels :

$$(\operatorname{ch} x)' = \frac{1}{2} (e^x + e^{-x})' = \frac{1}{2} (e^x - e^{-x}) = \operatorname{sh} x,$$

$$(\operatorname{sh} x)' = \frac{1}{2} (e^x - e^{-x})' = \frac{1}{2} (e^x + e^{-x}) = \operatorname{ch} x.$$

Nous voyons que, sous certains rapports, les fonctions hyperboliques sont plus simples que les fonctions trigonométriques. Le terme « fonctions hyperboliques » tient à ce que la courbe dans le plan (x, y) , dont la représentation paramétrique est

$$x = \operatorname{ch} t, \quad y = \operatorname{sh} t,$$

est l'hyperbole $x^2 - y^2 = 1$. Pour les fonctions trigonométriques, la courbe correspondante

$$x = \cos t, \quad y = \sin t$$

est la circonférence $x^2 + y^2 = 1$, c'est pourquoi les fonctions trigonométriques sont également appelées *fonctions circulaires*.

Exercices

1. Montrer que la fonction e^{-1/x^2} possède sur l'intervalle $[-a, a]$ les dérivées de tous les ordres, mais ne peut être représentée par une série de Taylor suivant les puissances de x .

2. Montrer que la fonction $y = x^{2n} \sin \frac{1}{x}$ possède, pour $x=0$, les dérivées successives, y compris la n -ième, mais que $y^{(n+1)}(0)$ n'existe pas.

3. Démontrer que si toutes les racines d'un polynôme $P(x)$ sont réelles, alors il en est de même de toutes les racines des polynômes $P'(x)$, $P''(x)$, ..., $P^{(n-1)}(x)$.

4. Supposons que $f(x)$ soit bornée et possède les dérivées $f'(x)$ et $f''(x)$ continues et bornées sur $(-\infty, \infty)$. Soit ($j = 0, 1, 2$) $M_j = \sup_{-\infty < x < \infty} |f^{(j)}(x)|$.
Montrer que

$$M_2 \leq 2M_0M_1.$$

5. Montrer que la dérivée seconde $f''(x)$ (si elle existe) peut être obtenue par un passage à la limite à partir des valeurs de la fonction $f(x)$ en trois points voisins, notamment :

$$f''(x) = \lim_{h \rightarrow 0} \frac{f(x) - 2f(x+h) + f(x+2h)}{h^2}.$$

Ecrire la formule analogue pour $f^{(n)}(x)$.

6. Si une fonction $f(x)$ convexe vers le bas (chapitre 7, exercice 7) a la dérivée seconde pour $x = x_0$, alors $f''(x_0) \geq 0$.

7. Si l'inégalité $f''(x) > 0$ est satisfaite partout sur $[a, b]$, alors la courbe $y = f(x)$ est convexe vers le bas sur $[a, b]$.

8. Démontrer que la fonction $w = \sin z$ n'a pas de racines dans le plan complexe autres que les racines réelles $0, \pm\pi, \pm2\pi, \dots$.

9. Démontrer que la fonction $w = \sin z$ est minorée en module par une constante positive sur l'ensemble des points de toutes les circonférences

$$|z| = \left(n + \frac{1}{2}\right)\pi, \quad n=0, 1, 2, \dots$$

10. (a) Evaluer l'erreur que l'on commet en remplaçant la fonction e^x dans l'intervalle $[0, 1]$ par son polynôme de Taylor du 10-ième degré. (b) Dans quel intervalle $[0, h]$ le polynôme de Taylor du 10-ième degré pour la fonction e^x donne la valeur de cette fonction à 10^{-7} près? (c) Quel doit être le degré du polynôme de Taylor de la fonction e^x pour qu'il permette de la calculer sur l'intervalle $[0, 1]$ à 10^{-7} près?

11. Dans un intervalle (a, b) contenant un point x_0 , la fonction $f(x)$ possède les dérivées successives, y compris la n -ième, et au point x_0 la dérivée d'ordre $n+1$. On sait que $f'(x_0) = \dots = f^{(n)}(x_0) = 0$, $f^{(n+1)}(x_0) \neq 0$. Décrire la position de la courbe $y = f(x)$ par rapport à sa tangente au point x_0 pour chacune des hypothèses suivantes :

$$(a) f^{(n+1)}(x_0) > 0, \quad n \text{ pair};$$

$$(b) f^{(n+1)}(x_0) > 0, \quad n \text{ impair};$$

$$(c) f^{(n+1)}(x_0) < 0, \quad n \text{ pair};$$

$$(d) f^{(n+1)}(x_0) < 0, \quad n \text{ impair}.$$

12. Si une fonction $f(x)$, au voisinage d'un point a , admet la représentation

$$f(a+h) = f(a) + Ah + o(h),$$

alors $f'(a)$ existe et vaut A (7.24). Supposons que la fonction $f(x)$, au voisinage du point a , admette la représentation

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}Bh^2 + o(h^2).$$

Est-ce que $f''(a)$ existe?

13. Démontrer l'inégalité

$$e - \sum_{k=0}^n \frac{1}{k!} < \frac{1}{n \cdot n!}.$$

14. En partant du fait que $n! \sum_{k=0}^n \frac{1}{k!}$ est un nombre entier démontrer que e est irrationnel.

15. Le résultat de l'exercice 12 chap. 6 tombe en défaut sans l'hypothèse que tous les ω_k sont de même signe. Montrer que, pour $\omega_k = (-1)^k \frac{1}{\sqrt{k}}$,

le produit $\prod_{k=1}^{\infty} (1 + \omega_k)$ diverge, bien que la série $\sum_{k=1}^{\infty} \omega_k$ converge. Mon-

trer que pour $\omega_k = (-1)^k \frac{1}{\sqrt{k}} - 1$ on a la situation inverse.

16. Montrer que, pour $-2\pi < x < 2\pi$, $n = 1, 2, \dots$, $x \neq 0$, on a les inégalités

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots - \frac{x^{4n-2}}{(4n-2)!} < \cos x < 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + \frac{x^{4n}}{(4n)!}.$$

17. En se servant des inégalités de l'exercice 16 pour $n = 1$, établir les évaluations suivantes pour le nombre π :

$$3,11 < \pi < 3,18.$$

Historique

Les dérivées d'ordre supérieur apparaissent en même temps que tout l'appareil du calcul différentiel et sont systématiquement appliquées par Newton et Leibniz (fin du XVII^e siècle) à des problèmes de géométrie, mécanique et physique. En ce qui concerne les applications aux problèmes mécaniques, l'interprétation selon laquelle l'accélération d'un mouvement est la dérivée seconde du chemin parcouru par rapport au temps est essentielle; la loi fondamentale de la mécanique (deuxième loi de Newton) lie la force qui agit sur un corps avec l'accélération et la masse du corps.

Les séries pour des fonctions élémentaires sont étudiées par Newton et James Gregory dès les années 1660. Au siècle suivant, Euler les prolonge dans le domaine complexe. L'œuvre de Taylor dans laquelle fait son apparition la « série de Taylor » se rapporte à 1715; d'ailleurs, Newton et Leibniz ont déjà réalisé les constructions équivalentes. Durant tout le XVIII^e siècle, les mathématiciens affirment ou sous-entendent que « toute fonction » se développe en sa série de Taylor; Cauchy est le premier à établir les conditions rigoureuses de convergence d'une série de Taylor vers une fonction donnée et à faire une nette distinction entre la convergence de cette série et sa convergence vers la fonction donnée (1823).

Intégrale de Riemann

L'aire d'une courbe est une quantité qui se produit sans cesse et s'augmente d'une fluxion continue, proportionnelle à l'ordonnée de la courbe.

Isaac Newton (1671)

L'aire au-dessous d'une courbe est mesurée par l'ordonnée de la quadratrice qui est la courbe dont le coefficient angulaire est l'ordonnée de la courbe donnée.

Leibniz (1693)

§ 9.1. Définition de l'intégrale et théorèmes d'existence

Tout comme la dérivée, l'intégrale est une des plus importantes notions de l'analyse mathématique classique.

9.11. Partition à points marqués. Si l'on a, sur un intervalle $a \leq x \leq b$, un système de points $x_0, x_1, \dots, x_n$ tel que

$$a = x_0 \leq x_1 \leq \dots \leq x_n = b,$$

alors on dit qu'il définit une *partition de l'intervalle* $[a, b]$ en intervalles $[x_k, x_{k+1}]$; les points $x_0, x_1, \dots, x_n$ s'appellent *points de division de l'intervalle* $[a, b]$. Si, de plus, on choisit, dans tout intervalle $[x_k, x_{k+1}]$, un point ξ_k

$$x_k \leq \xi_k \leq x_{k+1} \quad (k = 0, 1, \dots, n-1),$$

alors on dit qu'une *partition à points marqués* $\xi_0, \xi_1, \dots, \xi_{n-1}$ est donnée. Une partition est désignée par la lettre Π ou, sous une forme plus développée,

$$\Pi = [a = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq x_2 \leq \dots \leq \xi_{n-1} \leq x_n = b]. \quad (1)$$

Désignons la différence $x_{k+1} - x_k$ par Δx_k . Le nombre $d(\Pi) = \max \Delta x_k$ s'appelle *paramètre de la partition* Π .

9.12. *Somme intégrale.* Soit une fonction (finie) $y = f(x)$ donnée sur un intervalle $[a, b]$. Une partition (1) étant donnée, on peut former la « somme intégrale », c'est-à-dire le nombre

$$S_{\Pi}(f) = \sum_{k=0}^{n-1} f(\xi_k) \Delta x_k.$$

9.13. *Intégrale.* Un nombre I s'appelle *intégrale (de Riemann) d'une fonction $f(x)$ sur un intervalle fermé $[a, b]$* si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que, quelle que soit une partition Π avec $d(\Pi) < \delta$, l'inégalité

$$|I - S_{\Pi}(f)| < \varepsilon \quad (1)$$

a lieu.

Cette définition peut être considérée à la lumière de la définition générale d'une limite que l'on a donnée dans 4.12. Notamment, nous considérons l'ensemble E ayant pour éléments les partitions Π à points marqués. Pour un $\rho > 0$ donné, désignons par E_{ρ} la partie de E formée de toutes les partitions Π pour lesquelles $d(\Pi) < \rho$. Il est évident que deux ensembles quelconques E_{ρ_1} et E_{ρ_2} sont inclus l'un dans l'autre et que l'intersection de tous les E_{ρ} est vide. Les ensembles E_{ρ} définissent donc une direction S sur l'ensemble E dont on peut se servir afin de parler de la limite d'une fonction donnée sur E . Notamment, la définition 4.12 appliquée au cas en question est formulée comme suit :

Un nombre I est la limite de la fonction $S(\Pi)$ donnée sur l'ensemble des partitions Π à points marqués suivant la direction S si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $d(\Pi) < \delta$ implique l'inégalité

$$|I - S(\Pi)| < \varepsilon. \quad (2)$$

Une limite suivant la direction mentionnée S sur l'ensemble de toutes les partitions à points marqués sera aussi appelée *limite pour un morcellement illimité de la partition*.

Nous voyons maintenant que lorsque la fonction $S(\Pi)$ est construite d'après la fonction $f(x)$ de la manière indiquée dans 9.12, alors l'existence de la limite de $S(\Pi)$ suivant S , c'est-à-dire pour un morcellement illimité de la partition Π , est équivalente à l'existence de l'intégrale de la fonction $f(x)$. L'intégrale est désignée par le symbole

$$\int_a^b f(x) dx \quad (3)$$

qui ressemble à la somme intégrale initiale. Cependant, si le symbole de la somme intégrale renseigne sur les opérations qu'il faut effectuer pour la trouver, celui de l'intégrale représente, pour le moment, un tout « indécomposable », de sorte qu'il ne s'agit pas de multiplier

$f(x)$ par dx , puis « d'intégrer ». La raison pour laquelle on adopte ce symbole (et non pas $\int_a^b f(x)$, par exemple) sera mise en évidence dans la suite *). Or, il est évident que la désignation de la variable indépendante n'a aucune importance pour la valeur de l'intégrale, de sorte que, par exemple,

$$\int_a^b f(x) dx = \int_a^b f(\xi) d\xi = \int_a^b f(t) dt = \dots$$

Les nombres a et b s'appellent *limites d'intégration*; a est la limite inférieure et b la limite supérieure.

Pour une fonction $f(x)$, son intégrale n'existe pas forcément. Nous démontrerons plus bas (9.14 et 9.16) que toute fonction continue et certaines fonctions discontinues mais assez simples possèdent les intégrales; au contraire, des fonctions « très discontinues » (comme, par exemple, la fonction de Dirichlet (9.17)) n'ont pas d'intégrale. Une fonction $f(x)$ définie sur $[a, b]$ et dont l'intégrale sur $[a, b]$ existe est dite *intégrable* sur $[a, b]$.

9.14. A présent, nous allons démontrer l'existence de l'intégrale d'une fonction $y = f(x)$ continue sur un intervalle fermé $[a, b]$.

a. Désignons par $\omega_f(\delta)$ l'oscillation de la fonction $f(x)$ sur l'intervalle $[a, b]$: $\omega_f(\delta) = \sup_{|x' - x''| < \delta} |f(x') - f(x'')|$. Rappelons (5.17.c) que, pour une fonction $f(x)$ continue sur un intervalle $[a, b]$, on a toujours $\lim_{\delta \rightarrow 0} \omega_f(\delta) = 0$.

b. Nous dirons qu'une partition $\Pi' = [a = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq \dots \leq x_p = b]$ est *suiivante* par rapport à une partition $\Pi = [a = x_0 \leq \xi_0 \leq x_1 \leq \dots \leq x_n = b]$ si chacun des points $x_0, \dots, x_n$ figure parmi les points $x_0, \dots, x_p$ (de sorte que l'ensemble $x_0, \dots, x_p$ s'obtient en complétant l'ensemble $x_0, \dots, x_n$ par quelques nouveaux points de division); on n'impose aucune condition aux points marqués ξ_k, ξ'_k . Établissons maintenant deux lemmes simples.

c. **L e m m e.** Si une partition Π satisfait à la condition $d(\Pi) \leq \delta$, alors, pour toute partition *suiivante* Π' , on a :

$$|S_{\Pi'}(f) - S_{\Pi}(f)| \leq \omega_f(\delta)(b - a). \quad (1)$$

D é m o n s t r a t i o n. En passant de Π à Π' tout intervalle $[x_k, x_{k+1}]$ s'enrichit de nouveaux points de division et de nouveaux

*) « J'avertis qu'on prenne garde de ne pas omettre dx ... faute fréquemment commise, et qui empêche d'aller de l'avant, du fait qu'on ôte par là à ces indivisibles, comme ici dx , leur généralité... de laquelle naissent d'innombrables transfigurations et équipollences de figures » (Leibniz, 1686).

points marqués; en modifiant un peu les désignations on peut les noter comme suit:

$$x_k = x_{k0} \leq \xi_{k0} \leq x_{k1} \leq \xi_{k1} \leq x_{k2} \leq \dots \leq \xi_{km_k-1} \leq x_{km_k} = x_{k+1}.$$

La partie de la somme intégrale S_{Π} qui correspond à l'intervalle $[x_k, x_{k+1}]$ a la forme $\sum_{j=0}^{m_k-1} f(\xi_{kj}) \Delta x_{kj}$ ($\Delta x_{kj} = x_{k,j+1} - x_{kj}$).

Le terme $f(\xi_k) \Delta x_k$ de la somme intégrale S_{Π} peut être mis lui aussi sous la forme d'une somme $\sum_{j=0}^{m_k-1} f(\xi_k) \Delta x_{kj}$.

Nous avons

$$\sum_{j=0}^{m_k-1} f(\xi_{kj}) \Delta x_{kj} - \sum_{j=0}^{m_k-1} f(\xi_k) \Delta x_{kj} = \sum_{j=0}^{m_k-1} [f(\xi_{kj}) - f(\xi_k)] \Delta x_{kj},$$

et, comme $d(\Pi) < \delta$,

$$\begin{aligned} & \left| \sum_{j=0}^{m_k-1} f(\xi_{kj}) \Delta x_{kj} - \sum_{j=0}^{m_k-1} f(\xi_k) \Delta x_{kj} \right| \leq \\ & \leq \sum_{j=0}^{m_k-1} |f(\xi_{kj}) - f(\xi_k)| \Delta x_{kj} \leq \omega_f(\delta) \Delta x_k. \end{aligned}$$

En sommant suivant l'indice k on obtient

$$\left| \sum_k \sum_j f(\xi_{kj}) \Delta x_{kj} - \sum_k f(\xi_k) \Delta x_k \right| \leq \omega_f(\delta) (b-a),$$

et le lemme est démontré.

d. L e m m e. Pour deux partitions (arbitraires) Π_1 et Π_2 avec $d(\Pi_1) < \delta$, $d(\Pi_2) < \delta$, on a

$$|S_{\Pi_1}(f) - S_{\Pi_2}(f)| \leq 2\omega_f(\delta) (b-a). \quad (2)$$

Pour le démontrer, construisons une nouvelle partition Π_3 qui a pour points de division tous les points de division de Π_1 et Π_2 , les points marqués étant choisis d'une façon quelconque. La partition Π_3 est suivante par rapport à Π_1 ainsi qu'à Π_2 . D'après le lemme c, on a

$$|S_{\Pi_1}(f) - S_{\Pi_3}(f)| \leq \omega_f(\delta) (b-a),$$

$$|S_{\Pi_2}(f) - S_{\Pi_3}(f)| \leq \omega_f(\delta) (b-a),$$

d'où (2).

e. Nous sommes maintenant en mesure de démontrer le théorème fondamental.

Théorème. Toute fonction $f(x)$ définie et continue sur un intervalle fermé $[a, b]$ est intégrable sur $[a, b]$.

Démonstration. Etant donné le critère de Cauchy d'existence d'une limite suivant une direction (4.19), il suffit de montrer que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que

$$|S_{\Pi_1}(f) - S_{\Pi_2}(f)| \leq \varepsilon$$

si $d(\Pi_1) \leq \delta$, $d(\Pi_2) \leq \delta$.

La fonction $f(x)$ étant continue sur $[a, b]$, on peut trouver, pour un $\varepsilon > 0$, un $\delta > 0$ tel que $\omega_f(\delta) \leq \frac{\varepsilon}{2(b-a)}$. Alors, quelles que soient deux partitions Π_1 et Π_2 avec $d(\Pi) \leq \delta$, l'inégalité

$$|S_{\Pi_1}(f) - S_{\Pi_2}(f)| \leq 2 \frac{\varepsilon}{2(b-a)} (b-a) = \varepsilon$$

a lieu d'après le lemme d. Le critère de Cauchy est donc satisfait, et le théorème sur l'existence de l'intégrale d'une fonction continue est démontré.

9.15. Nous aurons également affaire aux intégrales des fonctions d'une classe plus vaste. Pour le moment, nous donnons quelques propriétés générales de l'intégrale dans la seule hypothèse qu'elle existe, sans nous servir de propriétés particulières des fonctions à intégrer.

a. Si une fonction $f(x)$ est intégrable sur un intervalle $[a, b]$, et si k est une constante, alors la fonction $kf(x)$ est intégrable sur $[a, b]$, et

$$\int_a^b kf(x) dx = k \int_a^b f(x) dx. \quad (1)$$

En effet, pour toute partition Π ,

$$S_{\Pi}(kf) = \sum_{j=0}^{n-1} kf(\xi_j) \Delta x_j = k \sum_{j=0}^{n-1} f(\xi_j) \Delta x_j = k S_{\Pi}(f),$$

le second membre ayant une limite pour un morcellement illimité de la partition Π , qui vaut $k \int_a^b f(x) dx$; il en résulte que le premier membre a la même limite, ce qui établit l'égalité (1).

b. Si deux fonctions $f(x)$ et $g(x)$ sont intégrables sur un intervalle $[a, b]$, alors la fonction $f(x) + g(x)$ est elle aussi intégrable sur $[a, b]$, et

$$\int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx. \quad (2)$$

En effet, pour tout Π , on a :

$$\begin{aligned} S_{\Pi}(f+g) &= \sum_{k=0}^{n-1} [f(\xi_k) + g(\xi_k)] \Delta x_k = \\ &= \sum_{k=0}^{n-1} f(\xi_k) \Delta x_k + \sum_{k=0}^{n-1} g(\xi_k) \Delta x_k = S_{\Pi}(f) + S_{\Pi}(g); \end{aligned}$$

en passant à la limite pour un morcellement illimité de la partition Π on établit l'existence de l'intégrale de la fonction $f(x) + g(x)$ ainsi que l'égalité (2).

c. *Toute fonction intégrable $f(x)$ est bornée.*

Supposons qu'une fonction $f(x)$ ne soit pas bornée sur $[a, b]$. Considérons une partition quelconque $\Pi = \{a = x_0 \leq x_1 \leq \dots \leq x_{n-1} \leq x_n = b\}$. Dans l'un au moins des intervalles $[x_j, x_{j+1}]$, la fonction $f(x)$ n'est pas bornée non plus. Alors, en fixant les points marqués dans tous les autres intervalles et en faisant varier ξ_j dans l'intervalle où $f(x)$ n'est pas bornée, on peut obtenir les valeurs de la somme intégrale aussi grandes (en module) que l'on veut. Les sommes 9.12 ne peuvent donc pas avoir une limite.

d. *Intégration d'inégalités. Si deux fonctions $f(x)$ et $g(x)$ intégrables sur $[a, b]$ vérifient l'inégalité*

$$f(x) \leq g(x),$$

alors

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (3)$$

En effet, les sommes intégrales $S_{\Pi}(f)$ et $S_{\Pi}(g)$, pour une même partition Π , avec les mêmes points marqués ξ_j , satisfont à l'inégalité

$$S_{\Pi}(f) = \sum_{j=1}^n f(\xi_j) \Delta x_j \leq \sum_{j=1}^n g(\xi_j) \Delta x_j = S_{\Pi}(g),$$

d'où l'on obtient (3) en passant à la limite pour $d(\Pi) \rightarrow 0$.

e. En particulier, si une fonction $f(x)$ intégrable sur un intervalle $[a, b]$ y est majorée par un nombre M , alors

$$\int_a^b f(x) dx \leq \int_a^b M dx = M(b-a). \quad (4)$$

f. D'une manière analogue, si une fonction $f(x)$ intégrable sur un intervalle $[a, b]$ y est minorée par un nombre m , alors

$$m(b-a) \leq \int_a^b f(x) dx. \quad (5)$$

g. En particulier, si la fonction $f(x)$ est non négative, on peut poser $m = 0$, et l'intégrale est elle aussi non négative. Si la fonction $f(x)$ est non positive, on pose $M = 0$ dans (4), et l'intégrale est non positive.

h. Pour une fonction $f(x)$ bornée,

$$m \leq f(x) \leq M,$$

on obtient l'estimation double de l'intégrale:

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a). \quad (6)$$

La quantité

$$\frac{1}{b-a} \int_a^b f(x) dx$$

s'appelle *moyenne (intégrale)* de la fonction $f(x)$. Il résulte de (6) l'inégalité double suivante pour la moyenne (« théorème de la moyenne »):

$$m \leq \frac{1}{b-a} \int_a^b f(x) dx \leq M. \quad (7)$$

Si la fonction $f(x)$ est continue sur $[a, b]$, alors, en choisissant pour m et M ses bornes inférieure et supérieure respectivement et en appliquant le théorème de Bolzano 5.22, nous concluons qu'il existe un point $\xi \in [a, b]$ tel que

$$\frac{1}{b-a} \int_a^b f(x) dx = f(\xi).$$

Le nombre $f(\xi)$ est la *valeur moyenne* de la fonction $f(x)$ sur l'intervalle $[a, b]$.

1. Si les fonctions $f(x)$ et $|f(x)|$ sont intégrables *), alors en intégrant l'inégalité

$$-|f(x)| \leq f(x) \leq |f(x)|$$

on trouve

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx,$$

d'où

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

*) L'intégrabilité de la fonction $f(x)$ implique, en réalité, celle de la fonction $|f(x)|$ (voir l'exercice 7).

j. Si une fonction $f(x)$ est intégrable sur chacun des intervalles $[a, c]$ et $[c, b]$, alors elle l'est aussi sur l'intervalle $[a, b]$, et

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (8)$$

Pour le démontrer, choisissons une partition quelconque de l'intervalle $[a, b]$:

$$\Pi = [a = x_0 \leq x_1 \leq \dots \leq x_m \leq c \leq x_{m+1} \leq \dots \leq x_n = b].$$

La somme intégrale correspondante, à points marqués ξ_k ($k = 0, 1, \dots, n-1$), se transforme de la façon suivante

$$\begin{aligned} S_{\Pi}(f) &= \sum_{k=0}^{n-1} f(\xi_k) \Delta x_k = \\ &= \sum_{k=0}^{m-1} f(\xi_k) \Delta x_k + f(\xi_m) \Delta x_m + \sum_{k=m+1}^{n-1} f(\xi_k) \Delta x_k = \\ &= \sum_{k=0}^{m-1} f(\xi_k) \Delta x_k + f(c)(c - x_m) + f(c)(x_{m+1} - c) + \\ &\quad + \sum_{k=m+1}^{n-1} f(\xi_k) \Delta x_k + [f(\xi_m) - f(c)] \Delta x_m = \\ &= S_{\Pi_1}(f) + S_{\Pi_2}(f) + [f(\xi_m) - f(c)] \Delta x_m, \end{aligned} \quad (9)$$

Π_1 et Π_2 étant des partitions des intervalles $[a, c]$ et $[c, b]$ respectivement. Lors d'un morcellement illimité de la partition Π , les partitions Π_1 et Π_2 sont aussi morcelées d'une façon illimitée, et le dernier terme dans (9) tend vers zéro (puisque la fonction intégrable $f(x)$ est bornée d'après c , et la longueur de l'intervalle Δx_m tend vers zéro). Le second membre de (9) tend donc vers une limite qui vaut

$$\int_a^c f(x) dx + \int_c^b f(x) dx.$$

Le premier membre de (9), à son tour, a $\int_a^b f(x) dx$ pour limite, d'où (8).

k. Inversement, si une fonction $f(x)$ est intégrable sur un intervalle $[a, b]$, alors elle est intégrable sur chacun des intervalles $[a, c]$ et $[c, b]$, où $a < c < b$. Nous nous bornons à la démonstration pour le segment $[a, c]$; l'autre cas s'épuise d'une façon analogue. Toute partition

$$\Pi = [a = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq \dots \leq x_m = c]$$

de l'intervalle $[a, c]$ peut être complétée de façon à obtenir une partition $\overline{\Pi}$ de l'intervalle $[a, b]$:

$$\overline{\Pi} = [a = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq \dots \leq x_m \leq \xi_m \leq x_{m+1} \leq \dots \leq x_n = b].$$

Pour une autre partition

$$\Pi' = [a = x'_0 \leq \xi'_0 \leq x'_1 \leq \xi'_1 \leq \dots \leq x'_p = c],$$

choisissons le même complément, de sorte que

$$\overline{\Pi}' = [a = x'_0 \leq \xi'_0 \leq x'_1 \leq \dots \leq x'_p \leq \xi_m \leq x_{m+1} \leq \dots \leq x_n = b].$$

Comme $f(x)$ est intégrable sur $[a, b]$, on peut trouver, pour tout $\varepsilon > 0$ donné, un $\delta > 0$ tel que

$$|S_{\overline{\Pi}}(f) - S_{\overline{\Pi}'}(f)| < \varepsilon$$

dès que $d(\overline{\Pi}) < \delta$, $d(\overline{\Pi}') < \delta$. Or, il est clair que, dans le cas en question, on a

$$S_{\overline{\Pi}}(f) - S_{\overline{\Pi}'}(f) = S_{\Pi}(f) - S_{\Pi'}(f),$$

donc, quels que soient les Π et Π' avec $d(\Pi) < \delta$, $d(\Pi') < \delta$, on a

$$|S_{\Pi}(f) - S_{\Pi'}(f)| < \varepsilon.$$

De la sorte, le critère de Cauchy pour les sommes intégrales de la fonction $f(x)$ sur l'intervalle $[a, c]$ est satisfait relativement à la direction $d(\Pi) \rightarrow 0$. D'où l'intégrabilité de la fonction $f(x)$ sur l'intervalle $[a, c]$.

1. En appliquant plusieurs fois la proposition k et compte tenu de j, on aboutit au résultat suivant:

Lorsqu'une fonction $f(x)$ est intégrable sur un intervalle $[a, b]$ et que $a = c_0 < c_1 < \dots < c_n = b$, alors $f(x)$ est intégrable sur chacun des intervalles $[c_{h-1}, c_h]$, et

$$\int_a^b f(x) dx = \int_{c_0}^{c_1} f(x) dx + \dots + \int_{c_{n-1}}^{c_n} f(x) dx.$$

m. Signalons, enfin, une estimation utile de la différence entre une intégrale et sa somme intégrale.

Si une fonction $f(x)$ est intégrable sur $[a, b]$ et $d(\Pi) \leq \delta$, alors

$$\left| \int_a^b f(x) dx - S_{\Pi}(f) \right| \leq \omega_f(\delta) \cdot (b-a).$$

Cette inégalité se déduit de l'inégalité 9.14 (1) par le passage à la limite pour $d(\Pi') \rightarrow 0$.

9.16. A présent, nous allons prouver l'existence de l'intégrale de certaines fonctions discontinues. Il est vrai que la classe des

fonctions discontinues que nous allons considérer n'est pas vaste, mais elle est suffisante pour toutes les applications prochaines.

a. Supposons qu'une fonction $h(x)$ soit nulle pour $a < x < b$ et prenne des valeurs quelconques $h(a)$ et $h(b)$. Montrons que $h(x)$ est intégrable sur $[a, b]$ et que

$$\int_a^b h(x) dx = 0. \quad (1)$$

En effet, la somme intégrale de la fonction $h(x)$, pour toute Π , se réduit à la forme

$$h(\xi_0) \Delta x_0 + h(\xi_{n-1}) \Delta x_{n-1},$$

où $h(\xi_0)$ est 0 ou $h(a)$ et $h(\xi_{n-1})$ est 0 ou $h(b)$. Lorsque $d(\Pi) \rightarrow 0$, cette expression tend, évidemment, vers zéro, ce que nous cherchions.

b. Soit une fonction $f(x)$ continue pour $a \leq x \leq b$. Supposons que $f_1(x)$ se confonde avec $f(x)$ pour $a < x < b$ et prenne des valeurs quelconques $f_1(a)$ et $f_1(b)$. Alors la fonction $f_1(x)$ est intégrable sur $[a, b]$, et

$$\int_a^b f_1(x) dx = \int_a^b f(x) dx. \quad (2)$$

En effet, la fonction $h(x) = f_1(x) - f(x)$ vérifie les conditions de a; elle est donc intégrable, et l'égalité (1) a lieu. Par conséquent, en vertu de 9.15.b, la fonction $f_1(x) = f(x) + h(x)$ est intégrable elle aussi, et la relation (2) a lieu d'après 9.15.b et (1), ce qu'il fallait démontrer.

c. Une fonction $f(x)$ est dite *continue par morceaux* sur un intervalle $[a, b]$ s'il existe une partition

$$a = x_0 < x_1 < \dots < x_n = b$$

telle que $f(x)$ soit continue dans tout intervalle ouvert $x_h < x < x_{h+1}$ et possède des limites finies $f(x_0 + 0)$, $f(x_1 - 0)$, $f(x_1 + 0)$, $f(x_2 - 0)$, $\dots$, $f(x_n - 0)$ (ses valeurs aux points x_h étant quelconques). Montrons que toute fonction $f(x)$ continue par morceaux sur un intervalle $[a, b]$ est intégrable sur cet intervalle.

Désignons par $f_h(x)$ la fonction qui est définie sur l'intervalle fermé $x_h \leq x \leq x_{h+1}$, se confond à son intérieur avec la fonction $f(x)$ et prend à ses extrémités les valeurs $f(x_h + 0)$ et $f(x_{h+1} - 0)$ respectivement. La fonction $f_h(x)$ est continue sur $[x_h, x_{h+1}]$; en vertu de b, la fonction $f(x)$ est intégrable sur $[x_h, x_{h+1}]$ et, de plus,

$$\int_{x_h}^{x_{h+1}} f(x) dx = \int_{x_h}^{x_{h+1}} f_h(x) dx.$$

Soient maintenant $g(x)$ une fonction continue par morceaux et $c_0 = a < c_1 < \dots < c_p = b$ ses points de discontinuité. Alors, d'après c, on a

$$\int_a^b g(x) dx = \sum_{h=0}^{p-1} \int_{c_h}^{c_{h+1}} g(x) dx = 0,$$

d'où, chaque terme étant non négatif,

$$\int_{c_h}^{c_{h+1}} g(x) dx = 0 \quad (h = 0, \dots, p-1).$$

La fonction $g(x)$ est continue à l'intérieur d'un intervalle $[c_h, c_{h+1}]$, donc $g(x) = 0$ en tout point intérieur de cet intervalle. L'inégalité $g(x) > 0$ n'est possible qu'aux points $c_0, \dots, c_p$ dont le nombre est fini. Le théorème est démontré.

9.17. Nous avons établi l'intégrabilité d'une fonction continue par morceaux; une telle fonction n'a qu'un nombre fini de points de discontinuité *). Que peut-on dire sur l'intégrabilité d'une fonction possédant un nombre infini de points de discontinuité? Considérons la fonction de Dirichlet $\chi(x)$ qui vaut 0 en tout point irrationnel et 1 en tout point rationnel de l'intervalle $[0, 1]$. Pour toute Π , si les points ξ_h sont rationnels, on obtiendra la somme intégrale $\sum \chi(\xi_h) \Delta x_h$ qui vaut 1; et si les points ξ_h sont irrationnels (pour la même Π), la somme intégrale $\sum \chi(\xi_h) \Delta x_h$ vaudra 0. Donc, pour la fonction de Dirichlet, il n'existe pas de limite des sommes intégrales pour un morcellement illimité de $[0, 1]$, elle n'est donc pas intégrable.

Ainsi, la frontière qui sépare les fonctions intégrables et non intégrables passe quelque part dans le domaine des fonctions dont les points de discontinuité sont en nombre infini. Il existe un théorème qui dit qu'une fonction $f(x)$ est intégrable au sens de Riemann si, et seulement si, l'ensemble de tous ses points de discontinuité peut être recouvert par une famille finie ou dénombrable d'intervalles dont la somme des longueurs est inférieure à ε , quel que soit $\varepsilon > 0$ (exercice 9). Cependant, pour intégrer des fonctions discontinues, il vaut mieux utiliser un autre schéma d'intégration, plus général que celui de Riemann, à savoir l'intégrale de Lebesgue [14].

*) On peut démontrer que toute fonction bornée avec des points de discontinuité en nombre fini est intégrable. Voir l'exercice 9.

9.18. Intégrale de b à a , où $a \leq b$. Soit $f(x)$ intégrable sur $a \leq x \leq b$. On pose, par définition :

$$\int_b^a f(x) dx = - \int_a^b f(x) dx. \quad (1)$$

Ceci fait, la formule

$$\int_{\alpha}^{\gamma} f(x) dx = \int_{\alpha}^{\beta} f(x) dx + \int_{\beta}^{\gamma} f(x) dx \quad (2)$$

devient valable quels que soient $\alpha, \beta, \gamma \in [a, b]$, la position relative des points α, β, γ sur l'intervalle $[a, b]$ étant quelconque. En effet, pour $\alpha < \beta < \gamma$, la formule a été démontrée dans 9.15.j. Considérons une autre position des points α, β, γ , par exemple $\gamma < \beta < \alpha$. Alors on sait que

$$\int_{\gamma}^{\alpha} f(x) dx = \int_{\gamma}^{\beta} f(x) dx + \int_{\beta}^{\alpha} f(x) dx.$$

Ensuite, en appliquant (1) on trouve

$$\begin{aligned} \int_{\alpha}^{\gamma} f(x) dx &= - \int_{\gamma}^{\alpha} f(x) dx = - \int_{\gamma}^{\beta} f(x) dx - \int_{\beta}^{\alpha} f(x) dx = \\ &= \int_{\beta}^{\gamma} f(x) dx + \int_{\beta}^{\alpha} f(x) dx, \end{aligned}$$

ce qu'il nous fallait ; pour toute autre position de α, β, γ , le raisonnement est analogue.

Signalons une relation qui découle immédiatement de (1) :

$$\frac{1}{a-b} \int_b^a f(x) dx = - \frac{1}{a-b} \int_a^b f(x) dx = \frac{1}{b-a} \int_a^b f(x) dx. \quad (3)$$

En appliquant au second membre de (3) l'estimation 9.15 (7) on a :

$$\inf f(x) = m \leq \frac{1}{a-b} \int_b^a f(x) dx \leq M = \sup f(x)$$

de sorte que, pour l'estimation de la moyenne, il est sans importance si $a < b$ ou bien $b < a$.

§ 9.2. A quoi sert l'Intégrale ?

Nous donnons dans 9.24 et 9.28 deux exemples géométriques d'application de la notion d'intégrale en analyse mathématique en remettant un exposé détaillé au § 9.6.

9.21. Aire d'un trapèze curviligne. La notion d'aire pour des figures géométriques simples, par exemple limitées par des segments de droite, est connue de la géométrie élémentaire. Mais qu'est-ce que l'« aire » d'une figure limitée par des lignes autres que segments de droite? La géométrie élémentaire ne nous en dit rien *).

Il est naturel de partir de l'hypothèse axiomatique suivante: l'aire d'une figure Φ est un nombre qui n'est pas plus grand que l'aire S de toute figure élémentaire (par exemple, composée de polygones) contenant Φ , ni plus petit que l'aire s de toute figure élémentaire contenue dans Φ (fig. 9.1).

Comme $s \leq S$ (d'après les théorèmes de la géométrie élémentaire), on a (1.62.b):

$$\sup s \leq \inf S,$$

où la borne supérieure dans le premier membre s'obtient en parcourant toutes les figures élémentaires incluses dans Φ et la borne inférieure du second en parcourant toutes les figures élémentaires contenant Φ . Si $\sup s = \inf S$, alors la seule valeur possible de l'aire de la figure Φ est la valeur commune de $\sup s$ et $\inf S$. Appliquons cette considération d'ordre général au cas d'un trapèze curviligne, c'est-à-dire d'une figure qui est limitée en bas par un segment $[a, b]$ de l'axe des x , en haut par le graphique d'une fonction continue $y = f(x) \geq 0$, à gauche et à droite par les ordonnées de ce graphique pour $x = a$ et $x = b$ respectivement. Considérons une partition $\Pi = [a = x_0 \leq \xi_0 \leq x_1 \leq \dots \leq x_n = b]$ et choisissons le point $\xi_h \in [x_h, x_{h+1}]$ de manière que la fonction $f(x)$ y prenne sa valeur maximale (sur $[x_h, x_{h+1}]$) que nous désignons par M_h . La somme intégrale correspondante

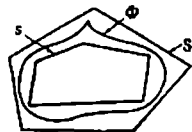


Fig. 9.1.

$$\bar{S}_{\Pi}(f) = \sum_{h=0}^{n-1} M_h \Delta x_h$$

peut être interprétée comme aire de la figure élémentaire formée par les rectangles exinscrits et contenant la figure Φ (fig. 9.2). Si l'on choisit le point $\xi_h \in [x_h, x_{h+1}]$ de manière que la fonction $f(x)$ y prenne sa valeur minimale (sur $[x_h, x_{h+1}]$) que nous désignons par m_h , alors la somme intégrale correspondante

$$\underline{S}_{\Pi}(f) = \sum_{h=0}^{n-1} m_h \Delta x_h$$

*) L'aire du cercle ne compte pas, puisqu'elle est calculée par des moyens non élémentaires (emploi de limites).

peut être interprétée comme aire de la figure élémentaire formée par les rectangles inscrits et contenue dans la figure (fig. 9.3). Donc, tout nombre S qui pourrait prétendre être l'« aire de la figure Φ » devrait bien, conformément à l'hypothèse, satisfaire à l'inégalité

$$S_{\Pi}(f) \leq S \leq \bar{S}_{\Pi}(f).$$

Or, si la fonction $f(x)$ possède l'intégrale (ce qui est démontré, pour une fonction continue, dans 9.14), alors, pour un morcellement

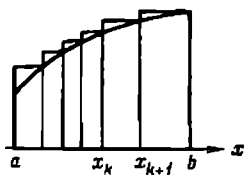


Fig. 9.2.

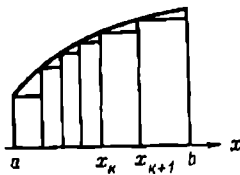


Fig. 9.3.

illimité de la partition Π , les sommes $S_{\Pi}(f)$ et $\bar{S}_{\Pi}(f)$ ont une limite commune I ; ainsi, nous arrivons à la seule définition possible de l'aire de la figure Φ :

$$S = I = \int_a^b f(x) dx.$$

9.22. Pour achever la construction, il faudrait établir que notre définition de l'aire d'un trapèze curviligne satisfait à l'hypothèse axiomatique formulée plus haut pour n'importe quelle figure élémentaire et non pas seulement pour celles qui figuraient dans la définition. On peut le faire, mais la démonstration est laborieuse et, de plus, le raisonnement n'est qu'un cas particulier des constructions générales de la théorie de la mesure d'aires et de volumes. Puisque nous n'en aurons besoin dans le présent livre, nous nous bornerons à un aperçu de cette théorie en renvoyant le lecteur pour les détails à la littérature *).

9.23. Pour commencer, donnons la définition de l'intégrale de Riemann sur un compact métrique. Soit K un compact (3.91). Appelons *cellules* certains sous-ensembles du compact K ; nous supposons que les propriétés suivantes soient valables:

- 1) le compact K lui-même et son sous-ensemble vide sont des cellules;
- 2) l'intersection de deux cellules est une cellule;
- 3) si une cellule A est contenue dans une cellule B , alors B peut être représentée comme réunion de cellules disjointes dont A ;
- 4) pour tout $\delta > 0$, il existe une partition du compact K en un nombre fini de cellules disjointes deux à deux dont les diamètres (3.12.s) ne dépassent pas δ .

Supposons ensuite qu'à toute cellule Q il corresponde un nombre non négatif mQ appelé *mesure de la cellule Q* , de sorte que chaque fois qu'une cellule Q

*) Voir, par exemple, [9] et [12].

est réunion des cellules disjointes $Q_1, \dots, Q_n$ la condition d'additivité:

$$mQ = mQ_1 + \dots + mQ_n \quad (1)$$

soit remplie.

Dans la suite, un compact muni d'un système de cellules et d'une mesure sera appelé *compact chargé*. On peut établir sur un compact chargé un schéma d'intégration analogue au schéma de l'intégrale de Riemann sur un intervalle fermé que l'on a décrit dans § 9.1.

Soit une fonction $f(x)$ définie sur un compact chargé K . Considérons une partition Π du compact K en un nombre fini de cellules disjointes $Q_1, \dots, Q_p$. Désignons par $d(\Pi)$ le plus grand des diamètres des cellules $Q_1, \dots, Q_p$. Choisissons, dans toute cellule Q_j , un point quelconque ξ_j et écrivons la somme intégrale

$$S_\Pi(f) = \sum_{j=1}^p f(\xi_j) mQ_j. \quad (2)$$

Un nombre I_f s'appelle *intégrale de Riemann de la fonction $f(x)$ sur le compact chargé K* si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que, quelle que soit une partition Π avec $d(\Pi) < \delta$, on ait l'inégalité

$$|I_f - S_\Pi(f)| < \varepsilon.$$

En d'autres termes, l'intégrale de Riemann de la fonction $f(x)$ est la limite des sommes intégrales (2) pour un morcellement illimité de la partition Π . Cette définition rentre dans le schéma général de limite suivant une direction, comme c'était le cas pour un intervalle fermé (9.13). On peut démontrer, de même que dans 9.14 pour le cas d'intervalle fermé, que toute fonction $f(x)$ continue est intégrable sur un compact chargé K .

9.24. Exemples.

a. Soit K un intervalle $[a, b]$ muni de sa métrique usuelle. Appelons cellules tous les intervalles $\alpha \leq x \leq \beta$, $\alpha < x \leq \beta$, $\alpha \leq x < \beta$, $\alpha < x < \beta$, où $\alpha \leq \beta$; la longueur de chaque intervalle sera la mesure de la cellule correspondante. Nous revenons alors à la définition usuelle de l'intégrale de Riemann sur l'intervalle fermé (9.13).

b. Soit K un pavé dans l'espace n -dimensionnel défini par les inégalités

$$a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n,$$

et muni de la métrique usuelle de l'espace euclidien (3.14). Appelons cellules les sous-pavés Q du pavé K définis par les inégalités $\alpha_1 \leq x_1 \leq \beta_1, \dots, \alpha_n \leq x_n \leq \beta_n$ ou bien par les inégalités qui s'en obtiennent en remplaçant n importe quels $\leq$ par $<$. Appelons mesure d'un pavé Q son volume, c'est-à-dire le

nombre $\prod_{h=1}^n (\beta_h - \alpha_h)$. Il est aisé de voir que toutes les conditions imposées à un système de cellules et à leur mesure sont satisfaites. L'intégrale ainsi obtenue s'appelle *intégrale de Riemann sur un pavé K n -dimensionnel*.

9.25. Nous dirons qu'une partie Z d'un compact chargé K a la mesure (jordanienne) nulle si, quel que soit $\varepsilon > 0$, elle peut être recouverte par un nombre fini de cellules $Q_1, \dots, Q_s$ avec $\sum_{j=1}^s mQ_j < \varepsilon$ et si, de plus, tout point $x \in Z$ est

un point intérieur (3.21) de l'ensemble $\bigcup_{j=1}^s Q_j$. Le théorème suivant (qui est une généralisation du théorème sur l'intégrabilité d'une fonction continue par morceaux 9.16) a lieu: toute fonction $f(x)$ bornée sur un compact chargé K et continue partout sauf sur un ensemble de mesure nulle est intégrable.

9.26. A tout ensemble $E \subset K$ faisons correspondre sa *fonction caractéristique* $\chi_E(x)$ qui vaut 1 sur E et 0 ailleurs. Nous dirons que l'ensemble E *possède un volume (mesure)* si la fonction $\chi_E(x)$ est intégrable auquel cas le nombre $mE = \int \chi_E$ sera appelé *volume (mesure de Jordan)* de l'ensemble E . Pour tout ensemble de mesure nulle, son volume est nul. Il résulte des propriétés fondamentales de l'intégrale que $E_1 \subset E_2$ implique $mE_1 \leq mE_2$ (si ces volumes existent) et que, étant donnée une partition d'un ensemble E possédant un volume en ensembles $E_1, \dots, E_s$ ayant aussi le volume, $mE = mE_1 + \dots + mE_s$ (additivité du volume).

Quels sont donc les ensembles $E \subset K$ possédant un volume? La question de l'existence du volume d'un ensemble E est tranchée en étudiant les propriétés de sa frontière. On appelle *frontière* d'un ensemble E l'ensemble des points $x \in K$ qui sont points limites pour E et pour son complémentaire à la fois (i.e. dont tout voisinage contient des points de E et ceux de $K - E$). Le théorème suivant a lieu: un ensemble $E \subset K$ a un volume si, et seulement si, la mesure de sa frontière est nulle. Dans la suite, de tels ensembles seront appelés *ensembles jordanens*. En particulier, pour toute cellule jordanienne, son volume est sa mesure initiale.

Pour les ensembles jordanens, la propriété d'additivité du volume peut être renforcée de la façon suivante: si un ensemble jordanien E est la réunion d'un nombre fini d'ensembles jordanens $E_1, \dots, E_s$ qui n'ont, deux à deux, aucun point intérieur commun, alors $mE = mE_1 + \dots + mE_s$.

9.27. Pour un pavé n -dimensionnel (9.24.b), on peut indiquer une classe simple mais assez riche de sous-ensembles jordanens. Notamment, tout ensemble borné dont la frontière est une réunion finie de surfaces d'équations de la forme

$$x_j = \varphi_j(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n),$$

où φ_j est une fonction continue (du point $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ par courant un domaine dans l'espace $(n-1)$ -dimensionnel), est bien jordanien.

En particulier, tout polyèdre (ensemble limité par un nombre fini de plans) est un ensemble jordanien.

Comme nous l'avons déjà dit dans 3.16.c, deux ensembles G et F dans l'espace euclidien n -dimensionnel R_n sont dits *géométriquement égaux* s'il existe une transformation isométrique de l'espace R_n dans lui-même qui transforme l'ensemble F en l'ensemble G . On démontre que les ensembles jordanens géométriquement égaux ont les volumes égaux. Ceci a lieu, en particulier, lorsque G s'obtient de F par une translation (3.15.b), une symétrie par rapport à un plan quelconque (3.15.a), une rotation (5.75). Ensuite, si un ensemble jordanien G s'obtient d'un ensemble jordanien F par une dilatation de rapport λ suivant une direction (2.67), alors $mG = \lambda mF$. Si un ensemble jordanien G s'obtient d'un ensemble jordanien F par une dilatation de rapport λ_1 le long de l'axe des x_1 , une

dilatation de rapport λ_2 le long de l'axe des x_2 , ..., une dilatation de rapport λ_n le long de l'axe des x_n , alors $mG = \lambda_1 \dots \lambda_n mF$. Enfin, si G s'obtient de F par une homothétie de rapport λ , alors $mG = \lambda^n mF$.

Dans le plan ($n = 2$), la figure décrite dans 9.21, i.e. le trapèze curviligne limité par les lignes $x = a$, $y = b$, $y = f(x) \geq 0$, est un ensemble jordanien, à condition que la fonction $f(x)$ soit continue, et, d'après le théorème général 9.26, elle possède un volume (que l'on doit, naturellement, appeler aire, pour $n = 2$). En vertu des propriétés du volume signalées dans 9.26, il est compris

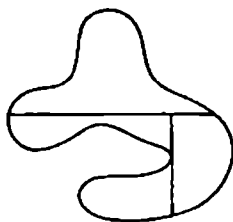


Figure partagée
en 4 trapèzes curvilignes

Fig. 9.4.

entre les volumes des figures élémentaires décrites dans 9.21, donc égal à l'expression que l'on y a donnée, c'est-à-dire à l'intégrale de Riemann de la fonction $f(x)$ sur l'intervalle $[a, b]$.

Ceci dit, l'aire d'une figure composée d'un nombre fini de trapèzes curvilignes (ou bien de figures géométriquement égales à de tels trapèzes) (fig. 9.4) peut être elle aussi exprimée par des intégrales de Riemann.

Ainsi, la théorie de la mesure d'aires dans le plan se réduit dans une grande mesure au calcul d'intégrales de Riemann. Cependant, ce n'est pas à toutes les questions naturelles que cette théorie peut répondre: en particulier, elle ne garantit l'existence du volume que pour une réunion finie d'ensembles possédant le volume. Une théorie plus générale, celle de Lebesgue, qui admet aussi bien les réunions dénombrables, est plus compliquée [14].

9.28. Longueur d'un arc de courbe. La notion de longueur d'un segment de droite et d'une ligne polygonale composée de segments nous est bien connue de la géométrie élémentaire. La longueur d'un arc de circonférence y est définie comme limite des longueurs de lignes polygonales inscrites et circonscrites. Sans insister sur la précision de cette définition, nous allons formuler directement une définition analogue pour un arc de courbe L quelconque que nous supposons donnée par une fonction

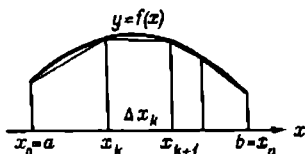


Fig. 9.5.

$y = f(x)$ sur un intervalle $[a, b]$. Soit une partition $\Pi = [a = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq x_2 \leq \dots \leq x_{n-1} \leq \xi_{n-1} \leq x_n = b]$ dont les points $\xi_0, \dots, \xi_{n-1}$ ne sont pas fixés pour le moment. Construisons une ligne polygonale L_Π aux sommets $(x_0, y_0), \dots, (x_n, y_n)$, où $y_k = f(x_k)$. Cette ligne est inscrite dans la courbe L (fig. 9.5). La longueur $s(L_\Pi)$ de la ligne L_Π peut être trouvée d'après la formule $(\Delta x_k = x_{k+1} - x_k, \Delta y_k = y_{k+1} - y_k)$:

$$s(L_\Pi) = \sum_{k=0}^{n-1} \sqrt{\Delta x_k^2 + \Delta y_k^2}.$$

Supposons que la fonction $f(x)$ ait la dérivée $f'(x)$.

Alors, en appliquant le théorème de Lagrange 7.44, on peut choisir un point $\xi_k \in (x_k, x_{k+1})$ de façon à avoir

$$\sqrt{\Delta x_k^2 + \Delta y_k^2} = \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x_k}\right)^2} \Delta x_k = \sqrt{1 + [f'(\xi_k)]^2} \Delta x_k.$$

L'expression de la longueur de L_Π se met alors sous la forme:

$$s(L_\Pi) = \sum_{k=0}^{n-1} \sqrt{1 + [f'(\xi_k)]^2} \Delta x_k. \quad (1)$$

C'est la somme intégrale pour la fonction $g(x) = \sqrt{1 + [f'(x)]^2}$ construite d'après la partition Π à points marqués $\xi_0, \xi_1, \dots, \xi_{n-1}$.

Si la fonction $g(x)$ est intégrable, les sommes (1) tendent, pour un morcellement illimité de la partition Π , vers la limite

$$\int_a^b \sqrt{1 + [f'(x)]^2} dx;$$

nous aboutissons ainsi à la définition suivante de la longueur $s(L)$ de l'arc L :

$$s(L) = \int_a^b \sqrt{1 + [f'(x)]^2} dx.$$

9.29. Sur le plan de calculs numériques, ce qu'on a dit ici ne vaudrait pas grand-chose et ne serait pas loin d'une tautologie, si la définition directe donnée dans 9.13 était le seul moyen de calculer les intégrales. En réalité, il existe les moyens efficaces du calcul d'intégrales qui sont exposés dans §§ 9.4 et 9.5.

Le § 9.3 est d'un caractère préliminaire; les résultats qu'on y obtient sont, d'ailleurs, importants en eux-mêmes.

§ 9.3. Intégrale comme fonction de sa limite supérieure

9.31. Soit $f(x)$ une fonction intégrable sur un intervalle $[a, b]$. L'intégrale

$$F(x) = \int_a^x f(t) dt$$

(on a désigné la variable sous le signe somme par t afin d'éviter la confusion avec la limite supérieure x) est une fonction du nombre x qui peut prendre toutes les valeurs de a à b . Vu la majoration

$$|F(x') - F(x'')| = \left| \int_{x''}^{x'} f(t) dt \right| \leq \sup_{x \in [a, b]} |f(x)| \cdot (x'' - x'),$$

la fonction $F(x)$ est continue.

Essayons de trouver la dérivée de cette fonction en un point quelconque $x = c$. D'après la propriété 9.18 (2), nous avons

$$\frac{F(c+h) - F(c)}{h} = \frac{1}{h} \left\{ \int_a^{c+h} f(t) dt - \int_a^c f(t) dt \right\} = \frac{1}{h} \int_c^{c+h} f(t) dt. \quad (1)$$

La valeur obtenue est la moyenne de la fonction $f(x)$ sur l'intervalle $[c, c+h]$ (*). Posons $m_h = \inf_{x \in [c, c+h]} f(x)$, $M_h = \sup_{x \in [c, c+h]} f(x)$;

*) Si $h < 0$, on pose par définition $[c, c+h] = \{x: c+h \leq x \leq c\}$.

alors, d'après l'inégalité 9.15 (7) (voir aussi 9.18 (3)), on a

$$m_h \leq \frac{1}{h} \int_c^{c+h} f(t) dt = \frac{F(c+h) - F(c)}{h} \leq M_h.$$

Si la fonction $f(x)$ est continue à droite en $x = c$, alors, pour $h \searrow 0$, les relations $m_h \rightarrow f(c)$ et $M_h \rightarrow f(c)$ ont lieu. De la sorte, la limite de l'expression (1) existe et vaut $f(c)$. Ceci veut dire que la fonction $F(x)$ possède, au point $x = c$, une dérivée à droite valant $f(c)$. D'une façon analogue, si $f(x)$ est continue à gauche pour $x = c$, l'expression (1) possède une limite $f(c)$ pour $h \nearrow 0$. En particulier, *en tout point $x = c$ de continuité de la fonction $f(x)$, la fonction $F(x)$ possède une dérivée qui vaut $f(c)$* . Pour $x = a$ et $x = b$, la fonction $F(x)$, en supposant toujours la continuité à droite ou à gauche respectivement de la fonction $f(x)$ en ces points, y possède la dérivée à gauche pour $x = b$ et la dérivée à droite pour $x = a$, qui sont respectivement $f(b)$ et $f(a)$.

9.32. Calcul d'intégrales à l'aide de fonctions primitives. Soit $y = f(x)$ une fonction continue par morceaux sur un intervalle $[a, b]$. Toute fonction continue $G(x)$ définie sur l'intervalle $[a, b]$ et possédant, en tout point de continuité x de la fonction $f(x)$, une dérivée qui vaut $f(x)$ est dite *primitive* de $f(x)$ sur $[a, b]$. Ainsi, la fonction x^n est une primitive de la fonction nx^{n-1} . Pour toute fonction $f(x)$ continue par morceaux, nous avons construit une primitive dans 9.31:

$$F(x) = \int_a^x f(t) dt.$$

Si $G(x)$ est une primitive de $f(x)$, toute fonction $G_1(x) = G(x) + C$ en est encore une; de la sorte, la fonction $f(x)$ possède une infinité de primitives qui ne diffèrent l'une de l'autre que par une constante. Réciproquement, *toute primitive $H(x)$ de la fonction $f(x)$ a la forme $G(x) + C$, où $G(x)$ est une primitive fixe*. En effet, en tout point de continuité de la fonction $f(x)$, nous avons $(H(x) - G(x))' = f(x) - f(x) = 0$. Par conséquent, $H(x) - G(x)$ est constant sur tout intervalle de continuité de $f(x)$ (7.45.c); $H(x) - G(x)$ étant continu sur tout $[a, b]$ de même que $H(x)$ et $G(x)$, $H(x) - G(x)$ est constant sur $[a, b]$.

Le fait signalé est à la base d'une méthode de calcul d'intégrales à l'aide de fonctions primitives. Supposons que l'on ait une primitive $G(x)$ de la fonction $f(x)$; alors n'importe quelle autre primitive, en particulier $F(x)$, diffère de $G(x)$ d'une constante, de sorte que

$$F(x) = \int_a^x f(t) dt = G(x) + C. \quad (2)$$

Pour trouver la constante C , posons $x=a$; alors, évidemment, $F(a)=0$, donc $C=-G(a)$. En posant $x=b$ dans (2) on trouve

$$\int_a^b f(t) dt = G(b) - G(a). \quad (3)$$

Ainsi, pour calculer l'intégrale d'une fonction $f(x)$ (continue par morceaux) sur un intervalle $[a, b]$, il faut trouver une primitive quelconque $G(x)$ de la fonction $f(x)$ et former la différence $G(b) - G(a)$.

La formule (3) obtenue en fait par Newton et Leibniz est l'une des plus importantes formules de l'analyse, qui lie le calcul différentiel avec le calcul intégral.

L'ensemble de toutes les fonctions primitives de $f(x)$ forme l'intégrale indéfinie de la fonction $f(x)$ et est désigné par

$$\int f(x) dx$$

(sans indication de limites); au contraire, l'intégrale 9.13 est appelée *intégrale définie*.

L'égalité (3) est encore notée comme suit:

$$\int_a^b f(x) dx = G(x) \Big|_{x=a}^{x=b} \equiv G(x) \Big|_a^b;$$

la notation dans les second et troisième membres indique qu'il faut poser $x=b$ puis $x=a$ dans $G(x)$ et retrancher ensuite le second nombre du premier.

Ainsi,

$$\int_0^b nx^{n-1} dx = x^n \Big|_0^b = b^n.$$

Ce résultat admet l'interprétation géométrique suivante: l'aire de la figure limitée par l'arc AC de la parabole du $(n-1)$ -ième degré $y = x^{n-1}$, le segment AB de l'axe des abscisses $0 \leq x \leq b$ et l'ordonnée BC au point $x=b$ fait $\frac{1}{n}$ -ième de l'aire du rectangle $ABCD$ (fig. 9.6).

9.33. d et $\int$ comme opérations inverses. Soient $f(x)$ une fonction définie et continue sur un intervalle $[a, b]$ et $F(x)$ une de ses primitives, de sorte que $F'(x) = f(x)$. Considérons les expressions $d \int f(x) dx$ et $\int dF(x)$.

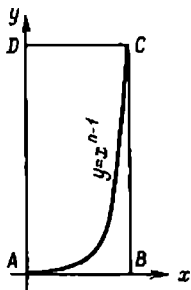


Fig. 9.6.

Selon la définition de l'intégrale indéfinie, on a

$$d \int f(x) dx = d [F(x) + C] = F'(x) dx = f(x) dx.$$

Ainsi, dans la combinaison $d \int$, les symboles d et $\int$ se réduisent. Or, d'après la définition de la différentielle et de l'intégrale indéfinie, on a

$$\int dF(x) = \int F'(x) dx = F(x) + C.$$

Ainsi, dans la combinaison $\int d$, les symboles $\int$ et d se réduisent, mais l'on doit ajouter une constante à $F(x)$.

9.34. Propriétés principales de l'intégrale indéfinie. En appliquant l'opération $\int$ aux règles connues de dérivation 7.33-7.34 et en se servant du résultat 9.33, on peut obtenir les règles correspondantes pour l'intégrale indéfinie. Il s'agit des règles suivantes de dérivation:

$$d(ku) = k du \quad (k \text{ constant}); \quad (1)$$

$$d(u + v) = du + dv; \quad (2)$$

$$d(uv) = u dv + v du. \quad (3)$$

Si $u = u(t)$, alors

$$dF(u) = F'(u) du = F'(u) u'(t) dt \quad (4)$$

(dérivation d'une fonction composée).

9.35. En appliquant l'opération $\int$ à l'égalité 9.34 (1) lue de droite à gauche on trouve, pour $k \neq 0$:

$$\int k du = \int d(ku) = ku + C = k(u + C_1) = k \int du. \quad (1)$$

9.36. En appliquant l'opération $\int$ à l'égalité 9.34 (2) on obtient

$$\int (du + dv) = \int d(u + v) = u + v + C = \int du + \int dv. \quad (1)$$

9.37. En appliquant l'opération $\int$ à 9.34 (3) et en se servant de 9.36 (1) on obtient

$$\int (u dv + v du) = \int u dv + \int v du = \int d(uv) = uv + C.$$

Mise sous la forme

$$\int u dv = uv - \int v du, \quad (1)$$

(C est incluse dans la deuxième intégrale), cette égalité s'appelle formule d'intégration par parties. Elle sert à trouver $\int u dv$, si $\int v du$ est donné.

9.38. La formule 9.34 (4) sera utilisée de la façon suivante. Soit $u = u(t)$ une fonction à dérivée continue, définie sur un intervalle $\alpha \leq t \leq \beta$ à valeurs dans un intervalle $A \leq u \leq B$; soit ensuite $f(u)$ une fonction continue, définie sur $A \leq u \leq B$. Considérons deux intégrales indéfinies:

$$F(u) = \int f(u) du \quad \text{et} \quad G(t) = \int f(u(t)) u'(t) dt.$$

La première est une fonction de u définie sur $[A, B]$, la seconde une fonction de t définie sur $[\alpha, \beta]$. Montrons que, pour $\alpha \leq t \leq \beta$, on a l'égalité

$$F[u(t)] = G(t) + C. \quad (1)$$

Il suffit de montrer que les différentielles du premier et du deuxième membre sont identiques. D'après la formule 9.34 (4), on a:

$$dF[u(t)] = F'(u) u'(t) dt = f(u(t)) u'(t) dt,$$

et, d'après la définition de l'intégrale indéfinie:

$$dG(t) = f(u(t)) u'(t) dt.$$

Il en résulte l'égalité (1) qui permet de formuler la règle d'intégration par changement de variable: dans une intégrale indéfinie $\int f(u) du$, on peut remplacer u par une fonction de t en remplaçant en même temps du par $u'(t) dt$ d'après la règle de transformation des différentielles. En d'autres termes, la propriété d'invariance d'une différentielle s'étend aux différentielles se trouvant sous le signe d'intégrale indéfinie.

Par la suite, on appliquera ces règles à de nombreux exemples.

9.39. Pour conclure ce paragraphe, nous démontrerons une proposition simple mais importante qui est la réciproque du théorème 9.31.

Une fonction $F(x)$ est dite lisse par morceaux sur un intervalle $[a, b]$ si elle est continue sur $[a, b]$ et possède, partout sauf en un nombre fini de points, une dérivée qui est continue par morceaux sur $[a, b]$.

En fait, nous avons établi dans 9.31 que l'intégrale

$$F(x) = \int_a^x f(t) dt$$

d'une fonction $f(x)$ continue par morceaux est une fonction lisse par morceaux. Réciproquement, soient $G(x)$ une fonction lisse par morceaux et $f(x)$ sa dérivée continue par morceaux. Puisque $G(x)$ est une primitive de $f(x)$ (9.32), l'égalité 9.32 (2) a lieu :

$$G(x) = \int_a^x f(t) dt + G(a).$$

Donc, toute fonction lisse par morceaux $G(x)$ est, à une constante additive près, l'intégrale d'une fonction continue par morceaux, avec x pour limite supérieure.

§ 9.4. Calcul des intégrales Indéfinies

9.41. Intégration des polynômes.

a. En appliquant l'opération $\int$ à la formule

$$d(x^n) = nx^{n-1} dx \quad (n = 1, 2, \dots)$$

on obtient, après avoir remplacé $n - 1$ par m et divisé par $n = m + 1$:

$$\int x^m dx = \frac{x^{m+1}}{m+1} + C \quad (m = 0, 1, 2, \dots),$$

où C est une constante arbitraire.

Il en résulte, d'après les règles 9.35-9.36, que, pour tout polynôme $P(x) = a_0 + a_1x + \dots + a_nx^n$, on a :

$$\int P(x) dx = a_0x + a_1 \frac{x^2}{2} + \dots + a_n \frac{x^{n+1}}{n+1} + C. \quad (1)$$

b. Dans certains cas, d'autres règles s'imposent. Par exemple, pour calculer l'intégrale

$$\int x(x^2 + 1)^{500} dx,$$

il n'est pas commode de développer le binôme $(x^2 + 1)^{500}$ d'après la formule de Newton et appliquer ensuite la formule générale (1); il vaut mieux faire la substitution $u = x^2 + 1$ et se servir de la règle 9.38 qui permet le calcul immédiat :

$$\int x(x^2 + 1)^{500} dx = \int u^{500} \frac{1}{2} du = \frac{1}{2} \cdot \frac{u^{501}}{501} + C = \frac{1}{1002} (x^2 + 1)^{501} + C.$$

Dans le cas de l'intégrale

$$\int x^3(x^2 + 1)^{500} dx,$$

c'est la règle d'intégration par parties 9.37 qui s'impose. On pose $u = x^2$, $dv = x(x^2 + 1)^{500} dx$; alors, d'après l'exemple précédent, $v = \frac{1}{1002} (x^3 + 1)^{501}$, $du = 2x dx$, donc

$$\begin{aligned} \int x^3 (x^2 + 1)^{500} dx &= uv - \int v du = \\ &= \frac{x^2}{1002} (x^3 + 1)^{501} - \int \frac{(x^2 + 1)^{501} \cdot 2x dx}{1002} = \\ &= \frac{x^2 (x^3 + 1)^{501}}{1002} - \frac{(x^3 + 1)^{502}}{1002 \cdot 502} + C. \end{aligned}$$

9.42. Intégration des fonctions rationnelles. On demande d'intégrer une fonction de la forme $f(x) = \frac{R(x)}{Q(x)}$, $R(x)$ et $Q(x)$ étant des polynômes à coefficients réels.

a. Si le degré du numérateur est au moins égal à celui du dénominateur, alors, en effectuant la division, on peut mettre $f(x)$ sous la forme

$$f(x) = P(x) + \frac{S(x)}{Q(x)},$$

où le degré de $S(x)$ est déjà inférieur à celui de $Q(x)$. Nous savons déjà (9.41) comment intégrer le polynôme $P(x)$, bornons-nous donc à la fraction régulière $\frac{S(x)}{Q(x)}$.

b. Soient $x_1, x_2, \dots, x_h$ les racines réelles du polynôme $Q(x)$, $z_1, \bar{z}_1, \dots, z_m, \bar{z}_m$ ses racines non réelles (en plus de toute racine $z = \alpha + i\beta$, la quantité complexe conjuguée $\bar{z} = \alpha - i\beta$ est, d'après 5.89.a, encore une racine de même multiplicité).

En vertu de 5.89.c, la fonction rationnelle $\frac{S(x)}{Q(x)}$ peut être décomposée en éléments simples :

$$\begin{aligned} \frac{S(x)}{Q(x)} &= \frac{A_{11}}{x - x_1} + \frac{A_{12}}{(x - x_1)^2} + \dots + \frac{A_{1r_1}}{(x - x_1)^{r_1}} + \frac{A_{21}}{x - x_2} + \dots + \\ &+ \frac{B_{11} + C_{11}x}{(x - \alpha_1)^2 + \beta_1^2} + \dots + \frac{B_{1q_1} + C_{1q_1}x}{[(x - \alpha_1)^2 + \beta_1^2]^{q_1}} + \dots + \\ &+ \frac{B_{m1} + C_{m1}x}{(x - \alpha_m)^2 + \beta_m^2} + \dots + \frac{B_{mq_m} + C_{mq_m}x}{[(x - \alpha_m)^2 + \beta_m^2]^{q_m}}, \end{aligned}$$

et le problème se réduit à l'intégration de ces éléments simples.

c. Pour $p \neq 1$, l'intégrale

$$\int \frac{dx}{(x - x_j)^p}$$

peut être trouvée à partir de l'égalité

$$d(x - x_j)^{1-p} = (1-p)(x - x_j)^{-p} dx;$$

en intégrant ses deux membres on obtient

$$\int \frac{dx}{(x - x_j)^p} = \frac{1}{1-p} (x - x_j)^{1-p} + C = \frac{1}{(1-p)(x - x_j)^{p-1}} + C. \quad (1)$$

d. Pour $p = 1$, on se sert de l'égalité 7.17 (3) qui, après une multiplication par dx , donne

$$d(\ln|x - x_j|) = \frac{dx}{x - x_j}.$$

d'où

$$\int \frac{dx}{x - x_j} = \ln|x - x_j| + C. \quad (2)$$

Notons que les formules (1), (2) ont lieu dans l'intervalle $[a, b]$ qui ne contient pas le point x_j (sinon, la fonction sous le signe somme ne serait pas continue par morceaux et 9.32 ne serait pas applicable).

e. L'intégrale

$$\int \frac{Ax + B}{[(x - \alpha)^2 + \beta^2]^m} dx$$

se ramène, par la substitution $x - \alpha = \beta u$, à une forme plus simple :

$$\int \frac{Cu + D}{(u^2 + 1)^m} du.$$

Si $D = 0$, en mettant $u^2 + 1 = v$ on est conduit au résultat final :

$$\begin{aligned} \int \frac{Cu du}{(u^2 + 1)^m} &= \frac{C}{2} \int \frac{dv}{v^m} = \\ &= \begin{cases} \frac{C}{2(1-m)} v^{m-1} + C_1 = \frac{C}{2(1-m)(u^2 + 1)^{m-1}} + C_1, & m > 1; \\ \frac{C}{2} \ln|v| + C_1 = \frac{C}{2} \ln(u^2 + 1) + C_1, & m = 1. \end{cases} \end{aligned}$$

f. Il reste l'intégrale

$$I_m = \int \frac{du}{(u^2 + 1)^m}. \quad (3)$$

Si $m = 1$, alors

$$\int \frac{du}{u^2 + 1} = \operatorname{arctg} u,$$

ce qui découle de la formule de dérivation 7.18.f. Soit $m > 1$. Appliquons à (3) la formule d'intégration par parties en posant

$\frac{1}{(u^2+1)^m} = v$, de sorte que $dv = -\frac{2mu \, du}{(u^2+1)^{m+1}}$. Nous obtenons

$$\begin{aligned} I_m &= \int \frac{du}{(u^2+1)^m} = \frac{u}{(u^2+1)^m} + 2m \int \frac{u \cdot u \, du}{(u^2+1)^{m+1}} = \\ &= \frac{u}{(u^2+1)^m} + 2m \left\{ \int \frac{(u^2+1) \, du}{(u^2+1)^{m+1}} - \int \frac{du}{(u^2+1)^{m+1}} \right\} = \\ &= \frac{u}{(u^2+1)^m} + 2m (I_m - I_{m+1}), \end{aligned}$$

d'où

$$I_{m+1} = \frac{1}{2m} \frac{u}{(u^2+1)^m} + \frac{2m-1}{2m} I_m.$$

Cette formule exprime I_{m+1} en fonction de I_m et une fonction rationnelle. En s'en servant pour $m=1$ on trouve $I_2 = \int \frac{du}{(u^2+1)^2} = \frac{1}{2} \frac{u}{(u^2+1)^2} + \frac{1}{2} \arctg u + C$. Posant $m=2$ on trouve la valeur de I_3 , etc.

Définitivement, toute fonction rationnelle possède une intégrale indéfinie qui s'exprime par fonctions rationnelles, logarithmes et arclangentes.

En pratique, il existe de nombreux procédés auxiliaires d'intégration des fonctions rationnelles que nous n'exposons pas [8].

9.43. Intégrales se ramenant à des intégrales de fonctions rationnelles. Certains types d'intégrales se ramènent à des intégrales des fonctions rationnelles par certaines substitutions standard. Dans la suite, R désigne une fonction rationnelle d'une ou deux variables.

a. L'intégrale

$$\int R(e^x) \, dx$$

devient rationnelle par la substitution $e^x = u$, $x = \ln u$, $dx = \frac{du}{u}$.

b. L'intégrale

$$\int R(\operatorname{tg} x) \, dx$$

devient rationnelle par la substitution $\operatorname{tg} x = u$, $x = \operatorname{arctg} u$, $dx = \frac{du}{1+u^2}$. En particulier,

$$\cos^2 x = \frac{1}{1+\operatorname{tg}^2 x}; \quad \sin^2 x = 1 - \cos^2 x; \quad \cos 2x = \cos^2 x - \sin^2 x;$$

$$\sin 2x = 2 \sin x \cos x = 2 \operatorname{tg} x \cos^2 x$$

sont des fonctions rationnelles de $\operatorname{tg} x$. Par conséquent, l'intégrale

$$\int R(\sin 2x, \cos 2x) dx$$

devient rationnelle elle aussi par la substitution $\operatorname{tg} x = u$.

Notons que $\sin x$, $\cos x$ ne sont pas des fonctions rationnelles de $\operatorname{tg} x$ (en effet, $\operatorname{tg} x$ a π pour période, et $\sin x$ et $\cos x$ sont de période 2π). Mais elles sont rationnelles en $u = \operatorname{tg} \frac{x}{2}$:

$$\sin x = \frac{2 \operatorname{tg} \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}}; \quad \cos x = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}};$$

donc, vu que $dx = \frac{2 du}{1 + u^2}$, l'intégrale

$$\int R(\sin x, \cos x) dx$$

se calcule rationnellement à l'aide de la substitution $\operatorname{tg} \frac{x}{2} = u$.

9.44. Intégrales des fonctions irrationnelles. Considérons l'intégrale

$$\int R(x, y) dx, \quad (1)$$

où y est une fonction de x . Si x et y peuvent s'exprimer rationnellement en une même variable t , de sorte que $x = x(t)$, $y = y(t)$, alors la rationalisation de l'intégrale (1) est immédiate:

$$\int R(x, y) dx = \int R(x(t), y(t)) x'(t) dt.$$

a. Soit, par exemple, $y = \sqrt{ax+b}$; on peut alors choisir la substitution $ax+b = t^2$, $x = \frac{1}{a}(t^2-b)$, $y = t$:

$$\int R(x, \sqrt{ax+b}) dx = \int R\left(\frac{1}{a}(t^2-b), t\right) \frac{2t dt}{a}.$$

b. Dans le cas où $y = \sqrt{\frac{ax+b}{cx+d}}$, on applique le même procédé:

$$\frac{ax+b}{cx+d} = t^2, \quad y = t, \quad x = \frac{t^2 d - b}{a - ct^2}.$$

c. Mais pour $y = \sqrt{ax^2 + bx + c}$, la substitution $ax^2 + bx + c = t^2$ n'est plus valable, puisque x n'est pas une fonction rationnelle de t . Ici on va procéder autrement. En dégageant sous le radical un carré parfait $(x - \alpha)^2$ et en remplaçant $(x - \alpha)$ par βu avec un

β convenablement choisi, on met la racine carrée sous une des formes suivantes :

$$y_1 = \sqrt{1-u^2}; \quad y_2 = \sqrt{1+u^2}; \quad y_3 = \sqrt{u^2-1}.$$

d. Dans le premier cas, la substitution $u = \sin \varphi$ (ou bien $\cos \varphi = u$) ramène l'intégrale (1) à une forme déjà vue (9.43)

$$\int R(u, \sqrt{1-u^2}) du = \int R(\sin \varphi, \cos \varphi) \cos \varphi d\varphi.$$

Exemples :

$$\int \frac{du}{\sqrt{1-u^2}} = \int \frac{\cos \varphi d\varphi}{\cos \varphi} = \int d\varphi = \varphi = \arcsin u + C \quad (2)$$

(substitution $u = \sin \varphi$);

$$\int \frac{du}{\sqrt{1-u^2}} = - \int \frac{\sin \theta d\theta}{-\sin \theta} = \int d\theta = \theta = \arccos u + C^* \quad (3)$$

(substitution $u = \cos \theta$).

La divergence des résultats obtenus par des substitutions différentes ne doit pas nous inquiéter : selon la théorie générale, le résultat est défini à une constante additive près, et nous voyons, en effet, que

$$\arccos u + \frac{\pi}{2} = \arcsin u. \quad (5.67 (1)).$$

e. Pour $y_2 = \sqrt{1+u^2}$, c'est la substitution hyperbolique $u = \operatorname{sh} \theta$, $1+u^2 = 1+\operatorname{sh}^2 \theta = \operatorname{ch}^2 \theta$ qui joue un rôle analogue, de sorte que

$$\int R(u, \sqrt{1+u^2}) du = \int R(\operatorname{sh} \theta, \operatorname{ch} \theta) \operatorname{ch} \theta d\theta.$$

Exemple :

$$\int \frac{du}{\sqrt{1+u^2}} = \int \frac{\operatorname{ch} \theta d\theta}{\operatorname{ch} \theta} = \int d\theta = \theta = \operatorname{arsh} u + C \quad (4)$$

(substitution $u = \operatorname{sh} \theta$).

La fonction $\operatorname{arsh} u$ (fig. 9.7) est l'inverse de $\operatorname{sh} \theta^{**}$. (9.61.g nous révélera sa signification géométrique et la raison pour laquelle la désignation ar est adoptée : *area* veut dire « aire » en latin.) Cette fonction peut être exprimée par des fonctions que nous avons étu-

*) Dans le présent cas, le domaine de définition de la fonction $\cos \theta$ est l'intervalle $]-\pi, 0]$. On a $+\sqrt{1-u^2} = -\sin \theta$, et on se sert de la fonction inverse croissante $\arccos_c u$ (5.67). Avec la fonction décroissante $\arccos_d u$, on aboutirait à la formule

$$\int \frac{du}{\sqrt{1-u^2}} = -\arccos_d u + C.$$

***) Nous gardons les notations de l'auteur. (N.d.R.)

diées en résolvant l'équation

$$u = \operatorname{sh} \theta = \frac{e^{\theta} - e^{-\theta}}{2} = \frac{1}{2} \left(e^{\theta} - \frac{1}{e^{\theta}} \right)$$

par rapport à θ . Nous avons alors $e^{\theta} = u + \sqrt{1+u^2}$ (le signe moins est à rejeter car $e^{\theta} > 0$), d'où

$$\theta = \operatorname{arsh} u \equiv \ln(u + \sqrt{1+u^2}).$$

f. Enfin, pour $y_3 = \sqrt{u^2 - 1}$, on aura recours à la substitution hyperbolique $u = \operatorname{ch} \theta$, $u^2 - 1 = \operatorname{ch}^2 \theta - 1 = \operatorname{sh}^2 \theta$, de sorte que

$$\int R(u, \sqrt{u^2 - 1}) du = \int R(\operatorname{ch} \theta, \operatorname{sh} \theta) \operatorname{sh} \theta d\theta.$$

Exemple :

$$\int \frac{du}{\sqrt{u^2 - 1}} = \int \frac{\operatorname{sh} \theta d\theta}{\operatorname{sh} \theta} = \int d\theta = \theta = \operatorname{arch} u + C \quad (5)$$

(substitution $u = \operatorname{ch} \theta$).

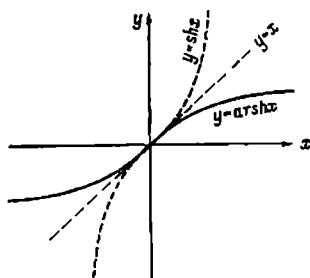


Fig. 9.7.

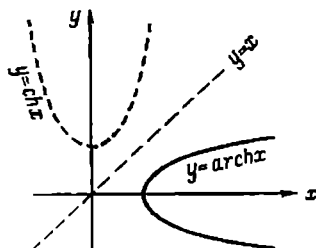


Fig. 9.8.

La fonction $\operatorname{arch} u$ (fig. 9.8) est l'inverse de $\operatorname{ch} \theta$. On peut l'expliquer de façon analogue en résolvant l'équation

$$u = \operatorname{ch} \theta = \frac{e^{\theta} + e^{-\theta}}{2} = \frac{1}{2} \left(e^{\theta} + \frac{1}{e^{\theta}} \right)$$

par rapport à θ . On a $e^{\theta} = u \pm \sqrt{u^2 - 1}$ (les deux signes conviennent), d'où

$$\theta = \operatorname{arch} u \equiv \ln(u \pm \sqrt{u^2 - 1}).$$

Les graphiques des fonctions $u = \operatorname{arsh} \theta$, $u = \operatorname{arch} \theta$ s'obtiennent de façon habituelle, i.e. sont symétriques des graphiques des fonc-

tions $u = \operatorname{sh} \theta$ et $u = \operatorname{ch} \theta$ respectivement par rapport à la bissectrice du premier et du troisième quadrant (5.39). Ils confirment que la fonction $\operatorname{arsh} \theta$ est univoque et $\operatorname{arch} \theta$ bivoque.

9.45. *Intégrales elliptiques.* Dans le cas où un polynôme $P(x)$ est de degré $n > 2$, l'intégrale

$$\int R(x, \sqrt{P(x)}) dx \quad (1)$$

ne s'exprime en général pas par des fonctions élémentaires *) (théorème d'Abel-Liouville; on en trouvera une démonstration dans [10] ou [11]). Pour $n = 3$ et $n = 4$, l'intégrale (1) est dite *elliptique*, pour $n > 4$ *hyperelliptique*. Il existe deux intégrales elliptiques les plus importantes (1-ère et 2-ième formes normales de Legendre):

$$\left. \begin{aligned} F(k, \varphi) &\equiv \int_0^{\sin \varphi} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \underset{(t=\sin \theta)}{=} \int_0^{\varphi} \frac{d\theta}{\sqrt{1-k^2 \sin^2 \theta}}, \\ E(k, \varphi) &\equiv \int_0^{\sin \varphi} \sqrt{\frac{1-k^2t^2}{1-t^2}} dt \underset{(t=\sin \theta)}{=} \int_0^{\varphi} \sqrt{1-k^2 \sin^2 \theta} d\theta. \end{aligned} \right\} 0 \leq k \leq 1$$

Pour $k = 0$, les deux intégrales représentent une même fonction:

$$F(0, \varphi) = \int_0^{\varphi} d\theta = \varphi, \quad E(0, \varphi) = \int_0^{\varphi} d\theta = \varphi.$$

Pour $k = 1$, les deux intégrales s'expriment par des fonctions élémentaires:

$$\begin{aligned} F(1, \varphi) &= \int_0^{\sin \varphi} \frac{dt}{1-t^2} = \frac{1}{2} \ln \frac{1+t}{1-t} \Big|_0^{\sin \varphi} = \frac{1}{2} \ln \frac{1+\sin \varphi}{1-\sin \varphi}; \\ E(1, \varphi) &= \int_0^{\sin \varphi} dt = t \Big|_0^{\sin \varphi} = \sin \varphi. \end{aligned}$$

Les tables des fonctions $F(k, \varphi)$ et $E(k, \varphi)$ sont à double entrée (pour les valeurs de la variable φ et du paramètre $k = \sin \alpha$). Ci-dessous une table abrégée de valeurs de ces fonctions:

*) C'est-à-dire celles que l'on obtient en appliquant à l'argument x les opérations algébriques (addition, soustraction, multiplication, division) en nombre fini, ainsi que les fonctions logarithmiques, trigonométriques et leurs inverses.

$\varphi \backslash \alpha$	$E(k, \varphi)$				$F(k, \varphi)$				
	90°	60°	45°	30°	0°	30°	45°	60°	90°
0°	0	0	0	0	0	0	0	0	0
30°	0,50	0,51	0,51	0,52	$0,52 = \frac{\pi}{6}$	0,53	0,54	0,54	0,55
45°	0,71	0,73	0,75	0,77	$0,78 = \frac{\pi}{4}$	0,80	0,82	0,85	0,88
60°	0,87	0,92	0,96	1,01	$1,05 = \frac{\pi}{3}$	1,09	1,14	1,21	1,32
90°	1	1,21	1,35	1,48	$1,57 = \frac{\pi}{2}$	1,69	1,85	2,16	

Les graphiques fig. 9.9 montrent la variation des fonctions $F(k, \varphi)$ et $E(k, \varphi)$ pour différentes valeurs de k .

Les fonctions elliptiques apparaissent lors du calcul de la longueur d'un arc d'ellipse (9.63.c), d'où leur nom. Elles trouvent des applications dans de nombreux problèmes d'analyse et sont étudiées tout aussi bien que les fonctions circulaires.

On peut montrer que l'intégrale elliptique générale

$$\int R(x, \sqrt{ax^4 + bx^3 + cx^2 + ex + f}) dx$$

se ramène par des transformations élémentaires aux 1-ère et 2-ième formes normales de Legendre, à des fonctions élémentaires et, éventuellement, à la

« forme normale de 3-ième espèce » à deux paramètres k et k' :

$$\int \frac{d\varphi}{(1 + h \sin^2 \varphi) \sqrt{1 - k^2 \sin^2 \varphi}}$$

(cf. [8]).

9.46. Quelques autres types d'intégrales.
a. Soit un polynôme $P(x)$. L'intégrale

$$\int P(x) e^x dx$$

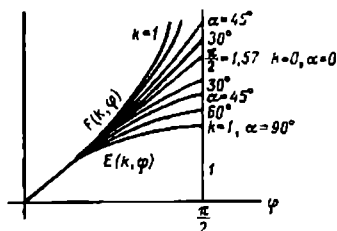


Fig. 9.9.

est une combinaison linéaire des intégrales de la forme

$$I_m = \int x^m e^x dx.$$

Une intégrale I_m se ramène par l'intégration par parties ($u = x^m$, $e^x dx = dv$; $du = mx^{m-1} dx$, $v = e^x$) à la forme

$$I_m = e^x x^m - m \int x^{m-1} e^x dx = e^x x^m - m I_{m-1}$$

et, puisque $I_0 = \int e^x dx = e^x + C$, s'exprime par des fonctions élémentaires.

b. Il en est de même des intégrales

$$\int P(x) \ln x dx, \quad \int P(x) \cos x dx, \quad \int P(x) \sin x dx.$$

c. Au contraire, les intégrales suivantes ne s'expriment pas par des fonctions élémentaires :

$$\int \frac{e^x}{x} dx \underset{(e^x=u)}{=} \int \frac{du}{\ln u} \quad (\text{logarithme intégral});$$

$$\int \frac{\sin x}{x} dx \quad (\text{sinus intégral});$$

$$\int \frac{\cos x}{x} dx \quad (\text{cosinus intégral}).$$

On trouvera les démonstrations des dernières propositions dans [18].

§ 9.5. Calcul des intégrales définies

9.51. Intégration par parties.

a. Soient u et v deux fonctions lisses par morceaux. Comme uv est une primitive de sa dérivée $uv' + vu'$, on a, d'après la formule de Newton-Leibniz 9.32 (2),

$$\int_a^b (uv' + vu') dx = \int_a^b (u dv + v du) = uv \Big|_{x=a}^b$$

ou bien

$$\int_a^b u dv = uv \Big|_{x=a}^b - \int_a^b v du.$$

Cette formule représente la règle d'intégration par parties pour une intégrale définie.

b. Exemple. Calculons l'intégrale $I_m = \int_0^{\pi/2} \sin^m x \, dx$. En intégrant par parties, trouvons

$$\int_0^{\pi/2} \sin^{m-1} x \, d(-\cos x) = -\sin^{m-1} x \cos x \Big|_0^{\pi/2} + (m-1) \int_0^{\pi/2} \sin^{m-2} x \cdot \cos^2 x \, dx.$$

A la suite de deux substitutions, le premier terme du second membre s'annule; on obtient donc

$$I_m = (m-1) \int_0^{\pi/2} \sin^{m-2} x (1 - \sin^2 x) \, dx = (m-1) I_{m-2} - (m-1) I_m,$$

d'où

$$I_m = \frac{m-1}{m} I_{m-2}.$$

Comme $I_0 = \int_0^{\pi/2} dx = \frac{\pi}{2}$, on a, pour $m = 2n$:

$$I_{2n} = \int_0^{\pi/2} \sin^{2n} x \, dx = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdot \dots \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{\pi}{2} = \frac{(2n-1)!!}{(2n)!!} \cdot \frac{\pi}{2}. \quad (1)$$

Comme $I_1 = \int_0^{\pi/2} \sin x \, dx = 1$, on a, pour $m = 2n+1$:

$$I_{2n+1} = \int_0^{\pi/2} \sin^{2n+1} x \, dx = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdot \dots \cdot \frac{4}{3} \cdot \frac{2}{1} \cdot 1 = \frac{(2n)!!}{(2n+1)!!}. \quad (2)$$

9.52.a. Formule de Taylor avec reste sous forme intégrale. Considérons d'abord l'intégrale

$$\int_a^b f(x) (b-x)^n \, dx.$$

En supposant que la fonction $f(x)$ soit lisse par morceaux intégrons par parties:

$$\begin{aligned} \int_a^b f(x) (b-x)^n \, dx &= -f(x) \frac{(b-x)^{n+1}}{n+1} \Big|_a^b + \int_a^b f'(x) \frac{(b-x)^{n+1}}{n+1} \, dx = \\ &= f(a) \frac{(b-a)^{n+1}}{n+1} + \frac{1}{n+1} \int_a^b f'(x) (b-x)^{n+1} \, dx. \end{aligned} \quad (1)$$

En nous servant de l'égalité (1) pour la fonction $f(x)$ qui possède sur l'intervalle $[a, b]$ les dérivées continues, y compris la n -ième, et la $(n+1)$ -ième dérivée continue par morceaux, nous obtenons successivement

$$\begin{aligned} f(b) - f(a) &= \int_a^b f'(x) dx = f'(a)(b-a) + \\ &+ \int_a^b f''(x)(b-x) dx = f'(a)(b-a) + \frac{1}{2} f''(a)(b-a)^2 + \dots \\ &+ \frac{1}{2} \int_a^b f'''(x)(b-x)^2 dx = \dots = f'(a)(b-a) + \\ &+ \frac{1}{2} f''(a)(b-a)^2 + \dots + \frac{1}{n!} f^{(n)}(a)(b-a)^n + \\ &+ \frac{1}{n!} \int_a^b f^{(n+1)}(x)(b-x)^n dx. \end{aligned}$$

Il en résulte la formule de Taylor

$$\begin{aligned} f(b) &= f(a) + f'(a)(b-a) + \frac{f''(a)}{1 \cdot 2} (b-a)^2 + \dots \\ &\dots + \frac{f^{(n)}(a)}{n!} (b-a)^n + R_n, \end{aligned} \quad (2)$$

où

$$R_n = \frac{1}{n!} \int_a^b f^{(n+1)}(x)(b-x)^n dx \quad (3)$$

est le reste sous forme intégrale de Cauchy.

b. A partir des inégalités

$$\begin{aligned} R_n &\leq \frac{1}{n!} \max f^{(n+1)}(x) \cdot \int_a^b (b-x)^n dx = \frac{(b-a)^{n+1}}{(n+1)!} \max f^{(n+1)}(x), \\ R_n &\geq \frac{1}{n!} \min f^{(n+1)}(x) \cdot \int_a^b (b-x)^n dx = \frac{(b-a)^{n+1}}{(n+1)!} \min f^{(n+1)}(x), \end{aligned}$$

on peut aussi mettre le reste sous forme de Lagrange (8.23 (2))

$$R_n = \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c),$$

en supposant que la fonction $f^{(n+1)}(x)$ soit continue et qu'elle prenne, par conséquent (5.23), toute valeur comprise entre ses valeurs maximale et minimale en un point c .

c. A titre d'application, considérons le problème de convergence de la série de Taylor (8.51) de la fonction

$$f(x) = (1+x)^\alpha \quad (4)$$

vers cette fonction. Dans le présent cas, la formule de Taylor (2) a la forme (pour $a = 0$, $b = x$):

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 + \dots + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{n!} x^n + \\ + \frac{\alpha(\alpha-1) \dots (\alpha-n)}{n!} \int_0^x (1+t)^{\alpha-n-1} (x-t)^n dt.$$

Si $|x| < 1$ et t est compris entre 0 et x , on a

$$t \leq \frac{t}{x}, \quad 1 - \frac{t}{x} = \frac{x-t}{x} \leq 1-t.$$

Comme x et $x-t$ sont de même signe,

$$\left| \frac{x-t}{x} \right| = \frac{x-t}{x} \leq 1-t,$$

donc

$$\frac{|x-t|}{1-t} \leq |x|.$$

Il en résulte

$$\left| \int_0^x \frac{(x-t)^n}{(1+t)^n} (1+t)^{\alpha-1} dt \right| \leq |x|^n \int_0^x (1+t)^{\alpha-1} dt = C |x|^n,$$

où $C = C(x, \alpha)$. Par conséquent

$$|R_n| \leq C \frac{|\alpha(\alpha-1) \dots (\alpha-n)|}{n!} |x|^n.$$

Désignons le second membre par ω_n . On a

$$\frac{\omega_{n+1}}{\omega_n} = \frac{|\alpha-n-1|}{n+1} |x|, \quad \lim_{n \rightarrow \infty} \frac{\omega_{n+1}}{\omega_n} = |x|;$$

donc, si $|x| < 1$, ω_n tendent vers 0. Ainsi, le reste (3) tend vers 0 si $|x| < 1$. Par conséquent, pour $-1 < x < 1$, le développement suivant a lieu

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 + \dots + \\ + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{1 \cdot 2 \dots n} x^n + \dots \quad (5)$$

Si α n'est pas un nombre naturel (le second membre n'est pas un polynôme), alors la série à droite diverge pour $|x| > 1$ (le rapport d'un terme au précédent est supérieur à 1), donc le rayon de convergence de la série (5) vaut 1.

En remplaçant dans (5) la variable réelle x par $z = x + iy$ on obtient le prolongement analytique de la fonction (4) dans le cercle $|z| < 1$:

$$(1+z)^\alpha = 1 + \alpha z + \frac{\alpha(\alpha-1)}{1 \cdot 2} z^2 + \dots + \frac{\alpha(\alpha-1) \dots (\alpha-n+1)}{1 \cdot 2 \dots n} z^n + \dots \quad (6)$$

Par exemple, pour $\alpha = -1$, on aboutit à la formule connue de sommation d'une progression géométrique convergente (6.11.b):

$$\frac{1}{1+z} = 1 - z + z^2 - \dots + (-1)^n z^n + \dots$$

d. On peut mettre la série (6) sous la forme

$$(1+z)^\alpha = 1 + (-\alpha)(-z) + \frac{(-\alpha)(1-\alpha)}{1 \cdot 2} (-z)^2 + \dots + \frac{(-\alpha)(1-\alpha) \dots (n-1-\alpha)}{1 \cdot 2 \dots n} (-z)^n + \dots$$

En appliquant les résultats 6.66.b on obtient: si $1 > -\alpha + 1$, c'est-à-dire $\alpha > 0$, alors la série (6) converge pour $-z = 1$, c'est-à-dire pour $z = -1$. En vertu de 6.67, on obtient une relation intéressante:

$$1 - \alpha + \frac{\alpha(\alpha-1)}{1 \cdot 2} - \frac{\alpha(\alpha-1)(\alpha-2)}{1 \cdot 2 \cdot 3} + \dots = 0; \quad (7)$$

d'après ce qu'on a vu, elle est valable pour tous les $\alpha > 0$.

Si $1 > -\alpha$, c'est-à-dire $\alpha > -1$, alors la série (6) converge pour $-z = -1$, c'est-à-dire pour $z = 1$; dans ce cas, 6.67 conduit à la relation:

$$1 + \alpha + \frac{\alpha(\alpha-1)}{1 \cdot 2} + \frac{\alpha(\alpha-1)(\alpha-2)}{1 \cdot 2 \cdot 3} + \dots = 2^\alpha \quad (8)$$

qui est valable pour $\alpha > -1$.

Les égalités (7) et (8) étendent aux exposants non entiers les propriétés connues des nombres $C_n^k = \frac{n!}{k!(n-k)!}$ pour k et n naturels:

$$\begin{aligned} C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n &= 2^n, \\ C_n^0 - C_n^1 + C_n^2 - \dots + (-1)^n C_n^n &= 0. \end{aligned}$$

9.53. Pour $a > b$, la formule de Newton-Leibniz 9.32 (3) :

$$\int_b^a f(x) dx = F(a) - F(b) \quad (F'(x) = f(x))$$

reste valable.

En effet, d'après 9.18, on a :

$$\int_b^a f(x) dx = - \int_a^b f(x) dx = -[F(b) - F(a)] = F(a) - F(b).$$

9.54. Intégration par changement de variable. Soit donnée une intégrale $\int_a^b f(u) du$, $f(u)$ étant une fonction continue sur l'intervalle $[a, b]$. Soit $u(t)$ une fonction définie et continûment dérivable sur un intervalle $[\alpha, \beta]$ et telle que $u(\alpha) = a$, $u(\beta) = b$. Un changement formel de variable $u = u(t)$ amène à l'égalité *)

$$\int_{u=a}^b f(u) du = \int_{t=\alpha}^{\beta} f[u(t)] u'(t) dt. \quad (1)$$

Dans quelles conditions cette égalité est-elle juste? La croissance de la fonction $u(t)$ sur l'intervalle $[\alpha, \beta]$ est une condition naturelle. Mais il se trouve que la formule (1) a lieu dans un cas plus général où la fonction $u(t)$ n'est pas monotone et même dans le cas où les valeurs de $u(t)$ peuvent sortir de l'intervalle $[a, b]$ sur lequel la fonction $f(u)$ est définie, pourvu que la fonction $f(u)$ admette un prolongement continu sur le segment plus grand $[A, B] \supset [a, b]$ dans lequel la fonction $u(t)$ prend ses valeurs pour $t \in [\alpha, \beta]$. Notamment, le théorème suivant a lieu :

Théorème. Si la fonction $u(t)$ définie et continûment dérivable sur $[\alpha, \beta]$ prend ses valeurs dans un intervalle $[A, B] \supset [a, b]$ et si la fonction $f(u)$ est continue sur $[A, B]$ alors la formule (1) est valable.

Démonstration. Soient $F(u)$ une primitive de la fonction $f(u)$ sur $[A, B]$ et $G(t)$ une primitive de $f(u(t)) u'(t)$ sur $[\alpha, \beta]$. D'après 9.38, on a

$$F(u(t)) = G(t) + C.$$

*) C'est pour cette raison que Leibniz met en garde contre le rejet de dx dans la désignation de l'intégrale (voir le renvoi page 258). Il est probable que « transfiguration » de Leibniz veuille dire passage de $f(u)$ à $f(u(t))$.

Par conséquent

$$\begin{aligned} \int_a^b f(u) du &= F(b) - F(a) = F(u(\beta)) - F(u(\alpha)) = \\ &= G(\beta) - G(\alpha) = \int_{\alpha}^{\beta} f(u(t)) u'(t) dt, \end{aligned}$$

et le théorème est démontré. L'égalité (1) exprime la *règle de changement de variable dans une intégrale définie*; soulignons que les limites d'intégration changent avec la fonction sous le signe somme. La formule de Newton-Leibniz étant aussi valable pour $\alpha > \beta$ (9.53), le cas $\alpha > \beta$ n'est pas exclu. En appliquant la règle de changement de variable, il est sans importance si les limites d'intégration sont disposées dans l'ordre naturel ou non.

9.55. Conséquences immédiates et exemples.

a. *Intégrale d'une fonction translatée.* Soit $f(x)$ une fonction définie sur un intervalle $[a, b]$; alors sa « translation » $f(x - h)$ est définie sur l'intervalle $[a + h, b + h]$ et l'on a

$$\int_{a+h}^{b+h} f(x-h) dx = \int_a^b f(x) dx.$$

Pour le démontrer, il suffit de réaliser la substitution $x - h = u$.

b. *Montrons que l'intégrale d'une fonction périodique *) $f(u)$ a une même valeur sur tout intervalle dont la longueur est égale à la période T de la fonction $f(u)$.* En effet, trouvons un nombre mT (m entier) compris entre a et $a + T$; notons

$$\int_a^{a+T} f(u) du = \int_a^{mT} f(u) du + \int_{mT}^{a+T} f(u) du;$$

dans la première intégrale faisons la substitution $u - (m-1)T = t$, dans la deuxième $u - mT = t$; étant donné que $f(u) \equiv f(u + mT) \equiv f(u + (m-1)T)$ en vertu de la périodicité de la fonction $f(u)$, on a

$$\int_a^{a+T} f(u) du = \int_{a-(m-1)T}^T f(t) dt + \int_0^{a-(m-1)T} f(t) dt = \int_0^T f(t) dt,$$

ce qu'il fallait démontrer.

*) C'est-à-dire vérifiant la relation $f(x + T) \equiv f(x)$ (cf. 5.65).

c. Exemple. Montrons que, quels que soient un $\gamma > 0$ et un k naturel, on a l'égalité

$$\int_{-\pi}^{\pi} \sin^{2\gamma} kx \, dx = \int_{-\pi}^{\pi} \cos^{2\gamma} kx \, dx. \quad (1)$$

En effet, en nous servant de l'égalité $\sin \alpha = \cos \left(\frac{\pi}{2} - \alpha \right)$ et de la substitution $\frac{\pi}{2} - kx = ku$ trouvons

$$\int_{-\pi}^{\pi} \sin^{2\gamma} kx \, dx = \int_{-\pi}^{\pi} \cos^{2\gamma} \left(\frac{\pi}{2} - kx \right) dx = \int_{-\frac{\pi}{2h} - \pi}^{-\frac{\pi}{2h} + \pi} \cos^{2\gamma} u \, du;$$

en appliquant la proposition b nous obtenons (1).

Dans le cas où 2γ est égal à un entier m , la valeur de l'intégrale (1) a été trouvée dans 9.51.b. Si $\gamma = 1$, la valeur des intégrales

$$\int_{-\pi}^{\pi} \sin^2 kx \, dx = \int_{-\pi}^{\pi} \cos^2 kx \, dx = \pi$$

se déduit immédiatement de l'égalité

$$\int_{-\pi}^{\pi} (\sin^2 kx + \cos^2 kx) \, dx = \int_{-\pi}^{\pi} 1 \cdot dx = 2\pi.$$

d. Intégrale de la symétrique d'une fonction. L'égalité:

$$\int_0^b f(x) \, dx = \int_0^b f(b-x) \, dx$$

est juste. En effet, la substitution $u = b - x$ a pour résultat

$$\int_a^b f(b-x) \, dx = - \int_b^a f(u) \, du = \int_0^b f(u) \, du = \int_0^b f(x) \, dx.$$

9.56. Pour terminer les considérations d'ordre général, étudions une notion générale importante.

Intégrale de Stieltjes. Soient $f(x)$ et $g(x)$ deux fonctions dans un intervalle $[a, b]$. Considérons une partition $\Pi = \{a = x_0 < \dots < x_n = b\}$ de l'intervalle $[a, b]$ à points marqués $\xi_j \in [x_j, x_{j+1}]$; ensuite, comme d'habitude, posons $d(\Pi) = \max(x_{j+1} - x_j)$. On appelle *somme intégrale de Stieltjes* le

nombre

$$S_{\Pi}(f, g) = \sum_{j=0}^{n-1} f(\xi_j) [g(x_{j+1}) - g(x_j)]. \quad (1)$$

Un nombre I est appelé *intégrale de Stieltjes d'une fonction $f(x)$ selon la fonction $g(x)$ sur l'intervalle $[a, b]$* si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que l'inégalité $|I - S_{\Pi}(f, g)| < \varepsilon$ a lieu pour n'importe quel Π avec $d(\Pi) < \delta$.

Autrement dit, l'intégrale de Stieltjes est la limite des sommes intégrales (1) pour un morcellement illimité de la partition Π . L'intégrale de Stieltjes est désignée par le symbole

$$\int_a^b f(x) dg(x). \quad (2)$$

Par exemple, on a toujours

$$\int_a^b dg(x) = g(b) - g(a).$$

Dans ce livre, nous nous passerons de l'étude générale de l'intégrale de Stieltjes [14] et démontrerons seul le *théorème d'existence de cette intégrale* dans le plus simple cas où $f(x)$ est continue sur $[a, b]$ et $g(x)$ y possède une dérivée continue. On verra que l'intégrale de Stieltjes se réduit alors à l'intégrale usuelle.

Une fois la fonction $g(x)$ dérivable, son accroissement sur $[x_j, x_{j+1}]$ peut être noté à l'aide du théorème de Lagrange 7.44 :

$$\Delta g(x_j) = g(x_{j+1}) - g(x_j) = g'(c_j) \Delta x_j, \quad c_j \in (x_j, x_{j+1}).$$

Par conséquent

$$\begin{aligned} \sum_{j=0}^{n-1} f(\xi_j) \Delta g(x_j) &= \sum_{j=0}^{n-1} f(\xi_j) g'(c_j) \Delta x_j = \\ &= \sum_{j=0}^{n-1} f(c_j) g'(c_j) \Delta x_j + \sum_{j=0}^{n-1} [f(\xi_j) - f(c_j)] g'(c_j) \Delta x_j = \\ &= \sum_{j=0}^{n-1} f(c_j) g'(c_j) \Delta x_j + \sum_{j=0}^{n-1} \varepsilon_j g'(c_j) \Delta x_j, \end{aligned} \quad (3)$$

où aucun des nombres $|\varepsilon_j|$, pour une partition Π assez fine, ne dépasse le $\varepsilon > 0$ donné. Alors

$$\left| \sum_{j=0}^{n-1} \varepsilon_j g'(c_j) \Delta x_j \right| \leq \varepsilon \cdot \max_{a \leq x \leq b} |g'(x)| (b-a),$$

de sorte que le deuxième terme du dernier membre de (3) tend vers zéro pour un morcellement illimité de la partition Π . Le premier

terme dans (3) est une somme intégrale pour la fonction $f(x) g'(x)$ et tend donc, suivant la même direction, vers la limite

$$\int_a^b f(x) g'(x) dx. \quad (4)$$

On voit que, dans les suppositions formulées plus haut ($f(x)$ continue, $g(x)$ continûment dérivable), l'intégrale de Stieltjes existe et est égale à l'intégrale usuelle (4).

On peut obtenir une formule analogue à (4) en supposant que $f(x)$ soit continue par morceaux ($g'(x)$ étant continue) ou $g'(x)$ soit continue par morceaux ($f(x)$ étant continue). Si les fonctions $f(x)$ et $g'(x)$ sont continues par morceaux les deux à la fois, alors, l'intégrale de Stieltjes cesse d'exister dès que leurs points de discontinuité se confondent.

Dans le cas général, on établit aisément les propriétés suivantes de l'intégrale de Stieltjes, par analogie avec les propriétés correspondantes de l'intégrale de Riemann (α, β sont des nombres):

$$\begin{aligned} & \int_a^b [\alpha_1 f_1(x) + \alpha_2 f_2(x)] dg(x) = \\ &= \alpha_1 \int_a^b f_1(x) dg(x) + \alpha_2 \int_a^b f_2(x) dg(x); \end{aligned} \quad (5)$$

$$\begin{aligned} & \int_a^b f(x) d[\beta_1 g_1(x) + \beta_2 g_2(x)] = \\ &= \beta_1 \int_a^b f(x) dg_1(x) + \beta_2 \int_a^b f(x) dg_2(x); \end{aligned} \quad (6)$$

$$\int_a^b f(x) dg(x) = \int_a^c f(x) dg(x) + \int_c^b f(x) dg(x); \quad (7)$$

$$\int_a^b f(x) dg(x) = f(x) g(x) \Big|_a^b - \int_a^b g(x) df(x), \quad (8)$$

si les intégrales dans les seconds membres existent.

§ 9.6. Applications de l'Intégrale

9.61. Calcul des aires. Nous avons vu dans 9.21 que le calcul d'aires des figures assez simples dans le plan se réduit au calcul d'intégrales définies. Lorsqu'on a un trapèze curviligne

limité par un segment $[a, b]$ de l'axe des x , une courbe continue $y = f(x) \geq 0$ et les ordonnées $x = a$, $x = b$, son aire s'exprime par l'intégrale (fig. 9.10)

$$I = \int_a^b f(x) dx. \quad (1)$$

Si une figure Φ n'est pas un trapèze curviligne, mais peut être représentée comme réunion de trapèzes curvilignes sans points intérieurs communs, alors son aire est la somme des aires des trapèzes. En général, l'aire d'une figure représentant la réunion de quelques figures simples (jordanienues) sans points intérieurs communs est égale à la somme des aires desdites figures simples.

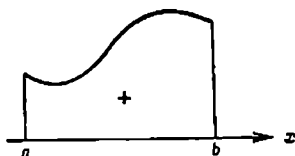


Fig. 9.10.

Revenons à la définition de l'aire sous la forme de l'intégrale (1). Elle est valable si $f(x) \geq 0$. Mais si $f(x) \leq 0$ (fig. 9.11), alors $I \leq 0$ et

I donne l'aire (limitée par un segment de l'axe, deux ordonnées et une courbe) précédée du signe moins. Si $f(x)$ change de signe dans $[a, b]$, alors l'intégrale I représente la somme algébrique



Fig. 9.11.

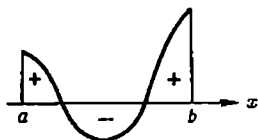


Fig. 9.12.

des aires (précédées du signe plus pour les parties de l'axe avec $f(x) \geq 0$ et du signe moins pour les parties avec $f(x) \leq 0$) (fig. 9.12).

Exemples.

a. Il résulte de 9.32 que l'aire limitée par la courbe $y = x^2$ et le segment $0 \leq x \leq b$ de l'axe fait un tiers de l'aire du rectangle correspondant (règle d'Archimède). A présent, déduisons une formule plus générale de Kepler (Nova stereometria solidorum vinariorum, 1615):

$$\int_a^b f(x) dx = \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \frac{b-a}{6}, \quad (2)$$

cette dernière étant valable si $f(x)$ est un polynôme du 3-ième degré au plus (fig. 9.13). En choisissant pour origine des coordonnées le milieu du segment

[a, b], on peut la mettre sous la forme

$$\int_{-c}^c f(x) dx = [f(-c) + 4f(0) + f(c)] \frac{c}{3}. \quad (3)$$

Notons maintenant que si la formule (3) est juste pour une fonction $f(x)$, elle l'est pour la fonction $kf(x)$ avec n'importe quel k ; si elle est juste pour deux fonctions $f(x)$ et $g(x)$, elle le reste pour la fonction $f(x) + g(x)$. Donc, afin d'établir cette formule pour toute fonction de la forme $f(x) = A + Bx + Cx^2 + Dx^3$, il suffit de la vérifier pour les fonctions $f_0(x) = 1$, $f_1(x) = x$, $f_2(x) = x^2$, $f_3(x) = x^3$. Pour $f_0(x) = 1$, elle est immédiate. Pour les fonctions $f_2(x)$ et $f_3(x)$ qui sont impaires, les deux membres de l'égalité (3) s'annulent, de sorte que la formule est toujours valable. Enfin, pour la fonction $f_1(x)$, la formule (3) se ramène à la règle d'Archimède; elle est donc établie pour tout polynôme d'un degré au plus égal à 3.

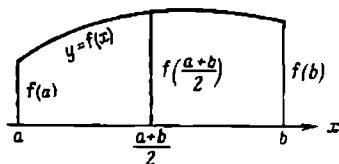


Fig. 9.13.

b. Aire du cercle $x^2 + y^2 \leq r^2$. Pour la raison de symétrie, on peut noter (fig. 9.14)

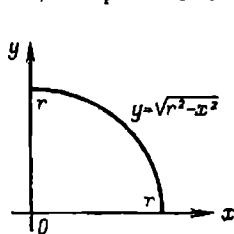


Fig. 9.14.

$$S = 4 \int_0^r y dx = 4 \int_0^r \sqrt{r^2 - x^2} dx.$$

En y posant $x = r \cos \theta$, $dx = -r \sin \theta d\theta$ on trouve :

$$\begin{aligned} S &= -4r^2 \int_0^{\pi/2} \sin^2 \theta d\theta = 4r^2 \int_0^{\pi/2} \sin^2 \theta d\theta = \\ &= 4r^2 \int_0^{\pi/2} \frac{1 - \cos 2\theta}{2} d\theta = 2r^2 \int_0^{\pi/2} d\theta - \\ &\quad - 2r^2 \int_0^{\pi/2} \cos 2\theta d\theta = \pi r^2 - r^2 \sin 2\theta \Big|_0^{\pi/2} = \pi r^2. \end{aligned}$$

Donc, le nombre $\pi/2$ que nous avons introduit plus haut comme première racine positive de la fonction $y = \cos x$ (5.65) trouve une nouvelle interprétation comme moitié de l'aire du cercle de rayon 1.

c. Pour l'aire du secteur du cercle $x^2 + y^2 \leq r^2$ (ou bien, sous forme paramétrique, $x = r \cos \theta$, $y = r \sin \theta$) limité par deux rayons $\theta = \theta_0$, $\theta = \theta_1$, $0 \leq \theta_0 < \theta_1 < \frac{\pi}{2}$ (fig. 9.15), on a

$$S(\theta_0, \theta_1) = S(0, \theta_1) - S(0, \theta_0),$$

de sorte qu'il suffit de calculer $S(0, \theta_0)$ pour tout $\theta_0 \in \left[0, \frac{\pi}{2}\right]$

(fig. 9.16). On a

$$\begin{aligned}
 S(0, \theta_0) &= \int_0^{r \cos \theta_0} x \operatorname{tg} \theta_0 dx + \int_{r \cos \theta_0}^r \sqrt{r^2 - x^2} dx = \\
 &= \frac{x^2}{2} \Big|_0^{r \cos \theta_0} \cdot \operatorname{tg} \theta_0 + r^2 \int_0^{\theta_0} \sin^2 \theta d\theta = \\
 &= \frac{r^2 \cos^2 \theta_0}{2} \operatorname{tg} \theta_0 + r^2 \int_0^{\theta_0} \frac{1 - \cos 2\theta}{2} d\theta = \\
 &= \frac{r^2}{2} \cos^2 \theta_0 \operatorname{tg} \theta_0 + \frac{r^2}{2} \theta_0 - \frac{r^2 \sin 2\theta}{4} \Big|_0^{\theta_0} = \frac{r^2}{2} \theta_0.
 \end{aligned}$$

Ainsi, on obtient

$$S(\theta_0, \theta_1) = \frac{r^2}{2} (\theta_1 - \theta_0). \quad (4)$$

On vérifie sans peine que la même formule reste valable pour tout secteur d'angle $0 \leq \theta_1 - \theta_0 < 2\pi$.

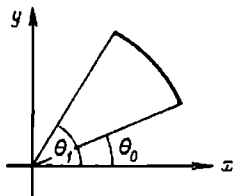


Fig. 9.15.

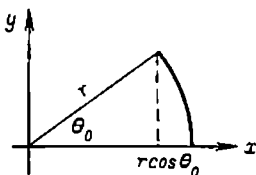


Fig. 9.16.

d. Lors d'une dilatation de rapport k suivant l'axe des y , l'aire S d'un trapèze curviligne (fig. 9.17) devient k fois plus grande. Cela résulte de l'égalité

$$S_h = \int_a^b k f(x) dx = k \int_a^b f(x) dx$$

(si l'on ne tient pas compte des considérations générales 9.27).

e. Lors d'une dilatation de rapport k suivant l'axe des x , l'aire S d'un trapèze curviligne (fig. 9.18) devient k fois plus grande. Pour le démontrer, faisons la substitution $x_1 = kx$:

$$S_h = \int_{ka}^{kh} f\left(\frac{x_1}{k}\right) dx_1 = k \int_a^b f(x) dx.$$

f. Aire d'une ellipse. L'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ s'obtient du cercle $x^2 + y^2 = 1$ à la suite d'une dilatation de rapport a le long de l'axe des x et d'une dilatation de rapport b le long de l'axe des y (fig. 9.19). Comme l'aire du cercle de rayon 1 vaut π , en appliquant les résultats des exemples d et e, on trouve $S_{\text{ellipse}} = \pi ab$.

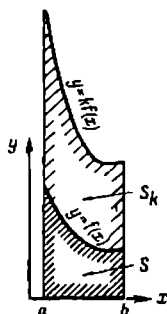


Fig. 9.17.

g. Interprétation géométrique des fonctions hyperboliques inverses. Considérons l'hyperbole $x^2 - y^2 = 1$

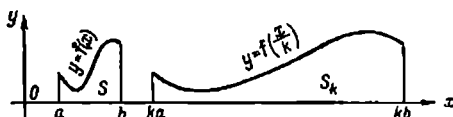


Fig. 9.18.

et trouvons l'aire du secteur ABC (fig. 9.20). Cette aire est donnée par l'intégrale

$$I = \int_1^x \sqrt{u^2 - 1} du$$

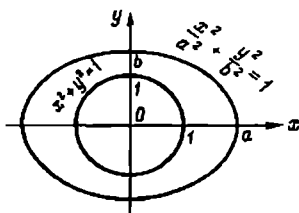


Fig. 9.19.

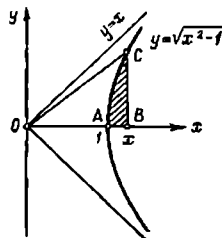


Fig. 9.20.

que nous allons calculer par la substitution $u = \cosh t$:

$$\int_1^x \sqrt{u^2 - 1} du = \int_0^{\operatorname{arsh} x} \operatorname{sh}^2 t dt.$$

Comme $\operatorname{ch} 2t = \operatorname{ch}^2 t + \operatorname{sh}^2 t$ et $1 = \operatorname{ch}^2 t - \operatorname{sh}^2 t$, on a $\operatorname{sh}^2 t = \frac{\operatorname{ch} 2t - 1}{2}$, d'où

$$I = \int_0^{\operatorname{arch} x} \frac{\operatorname{ch} 2t - 1}{2} dt = \left[\frac{1}{4} \operatorname{sh} 2t - \frac{t}{2} \right]_{t=0}^{\operatorname{arch} x}.$$

Notons que $\operatorname{sh} 2t = 2 \operatorname{sh} t \operatorname{ch} t$, donc $\operatorname{sh} 2t = 2 \sqrt{\operatorname{ch}^2 t - 1} \operatorname{ch} t$ et, par conséquent :

$$I = \frac{1}{2} [x \sqrt{x^2 - 1} - \operatorname{arch} x] = \frac{1}{2} x \sqrt{x^2 - 1} - \frac{1}{2} \operatorname{arch} x.$$

Le premier terme représente l'aire du triangle OBC ; ainsi, $\frac{1}{2} \operatorname{arch} x$ est l'aire du secteur OAC , donc $S = \operatorname{arch} x$ est l'aire du secteur $OC'AC$ (fig. 9.21). En inversant cette formule on obtient une représentation paramétrique de l'hyperbole :

$$x = \operatorname{ch} S, \quad y = \sqrt{x^2 - 1} = \operatorname{sh} S.$$

A vrai dire, nous avons déjà constaté que la courbe d'équations paramétriques $x = \operatorname{ch} t$, $y = \operatorname{sh} t$ est une hyperbole; mais ce n'est que maintenant que l'on a découvert la signification géométrique du paramètre t qui n'est rien autre que l'aire du secteur déterminé par un point donné de l'hyperbole.

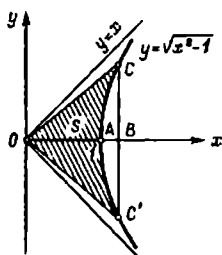


Fig. 9.21.

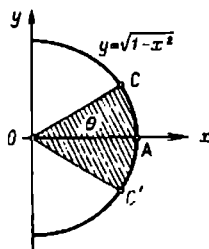


Fig. 9.22.

Il est utile de comparer cette représentation à la représentation paramétrique de la circonférence

$$x = \cos \theta, \quad y = \sin \theta.$$

Ici θ est l'angle polaire AOC ; or, cette même quantité représente l'aire du secteur $OC'AC$ (fig. 9.22). Ainsi, pour la circonférence et pour l'hyperbole, l'aire du secteur peut jouer le rôle de paramètre principal.

h. Inégalité de Young. Soit $y = f(x)$ ($0 \leq x < \infty$) une fonction croissante et continue, telle que $f(0) = 0$, $f(\infty) = \infty$. Soit $x = g(y)$ la fonction inverse (5.38) qui, dans le présent cas, est continue elle aussi et croît de $x = 0$ à $x = \infty$. Posons pour $x \geq 0$, $y \geq 0$:

$$F(x) = \int_0^x f(\xi) d\xi, \quad G(y) = \int_0^y g(\eta) d\eta.$$

Quels que soient $x > 0$, $y > 0$, l'inégalité suivante, dite de Young, a lieu:

$$xy \leq F(x) + G(y). \quad (5)$$

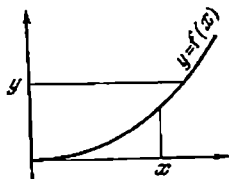


Fig. 9.23.

Elle est immédiate si l'on part des considérations géométriques (fig. 9.23); le lecteur construira sans doute une démonstration analytique. Notons encore que l'inégalité (5) devient égalité dans un seul cas où $y = x$.

L'inégalité de Young sert de point de départ pour de nombreuses autres inégalités importantes de l'analyse. Signalons ici quelques cas particuliers: en posant $f(x) = x^{p-1}$, $g(y) = y^{q-1}$ on obtient l'inégalité

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \left(\frac{1}{p} + \frac{1}{q} = 1 \right) \quad (6)$$

et si l'on pose $y = \ln(x+1)$, $x = e^y - 1$, on aboutit à l'inégalité

$$xy \leq (x+1) \ln(x+1) - x + e^y - y - 1. \quad (7)$$

Les deux inégalités (6) et (7) sont valables quels que soient $x \geq 0$, $y \geq 0$.

9.62. Aire en coordonnées polaires. Considérons une courbe en coordonnées polaires (fig. 9.24) $r = r(\varphi)$ et calculons l'aire du secteur OAB limité par deux rayons $\varphi = \alpha$ et $\varphi = \beta$ et par l'arc de la courbe $r = r(\varphi)$ qu'ils interceptent. Le secteur OAB est un ensemble jordanien et, selon la théorie générale (9.27), son aire est comprise entre l'aire de tout ensemble jordanien contenant le secteur OAB et celle de tout ensemble jordanien contenu dans ce secteur. Soit

$$\Pi = \{\alpha = \varphi_0 \leq \varphi_1 \leq \dots \leq \varphi_n = \beta\}$$

une partition de l'intervalle $[\alpha, \beta]$ de variation de la coordonnée φ , et soit, comme d'habitude,

$$\Delta_k \varphi = \varphi_{k+1} - \varphi_k, \quad m_k = \min_{\varphi_k \leq \varphi \leq \varphi_{k+1}} r(\varphi),$$

$$M_k = \max_{\varphi_k \leq \varphi \leq \varphi_{k+1}} r(\varphi).$$

Désignons par P_k un point $(\varphi_k, r(\varphi_k))$ de la courbe $r = r(\varphi)$. On connaît l'expression (9.61 (4)) pour l'aire d'un secteur limité par deux rayons $\theta = \varphi_k$ et

$\theta = \varphi_{k+1}$, et deux arcs des circonférences de rayons m_k et M_k (fig. 9.25). En se servant de ces expressions on remarque que la quantité S qui est à déterminer vérifie les inégalités

$$\frac{1}{2} \sum_{k=0}^{n-1} m_k \Delta \varphi_k \leq S \leq \frac{1}{2} \sum_{k=0}^{n-1} M_k \Delta \varphi_k.$$

Or, si la fonction $r = r(\varphi)$ est continue, alors, pour un morcellement illimité

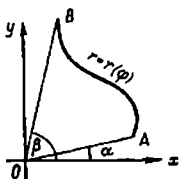


Fig. 9.24.

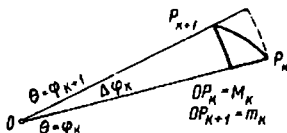


Fig. 9.25.

de la partition Π , les deux sommes tendent vers la même limite

$$\frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi;$$

ainsi, cette quantité représente l'aire de la figure OAB .

Exemple. La courbe $r = a \sin m\varphi$ (pour m impair) s'appelle rosace à m branches (fig. 9.26). La relation

$$r = a \sin 3\varphi \quad \frac{1}{2} \int_0^{\pi/m} a^2 \sin^2 m\varphi d\varphi = \frac{1}{2m} \int_0^{\pi} a^2 \sin^2 \psi d\psi$$

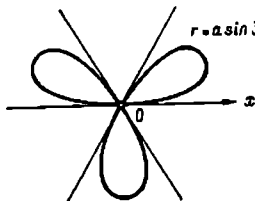


Fig. 9.26.

montre que l'aire d'une boucle de la rosace à m branches fait la $\left(\frac{1}{m}\right)^{\text{ième}}$ partie de l'aire de la « rosace à une branche » $r = a \sin \varphi$ (qui n'est, d'ailleurs, rien d'autre qu'une circonférence de diamètre a).

9.63. Longueur d'un arc. En définissant la longueur d'un arc comme limite des longueurs des lignes polygonales inscrites, on a obtenu (9.28), pour la longueur de l'arc de la courbe $y = f(x)$ limité par deux points $x = a$ et $x = b$, l'expression suivante:

$$s = \int_a^b \sqrt{1 + [f'(x)]^2} dx \quad (1)$$

(dans la supposition que l'intégrale existe).

Dans le cas général, on peut décrire une courbe γ par ses équations paramétriques qui représentent les coordonnées courantes d'un

point de la courbe comme fonctions d'un paramètre t . Ainsi, on aura deux équations dans le plan: $x = x(t)$, $y = y(t)$, et trois dans l'espace tridimensionnel: $x = x(t)$, $y = y(t)$, $z = z(t)$; dans l'espace n -dimensionnel, on aura n équations:

$$x_1 = x_1(t), x_2 = x_2(t), \dots, x_n = x_n(t). \quad (2)$$

Supposons que les fonctions $x_j(t)$ soient définies sur un intervalle $a \leq t \leq b$ et y possèdent les dérivées continues par morceaux. La courbe γ est alors dite *lisse par morceaux*.

Soit $\Pi = \{a = t_0 < t_1 < \dots < t_p = b\}$ une partition de l'intervalle $[a, b]$. A tout point t_i il correspond un point M_i de la courbe γ ayant $x_1(t_i)$, $\dots$, $x_n(t_i)$ pour coordonnées. En joignant les points M_i par les segments, on obtient une ligne polygonale $M_0 M_1 \dots M_p$ dont la longueur est donnée par la formule

$$s_p = \sum_{i=0}^{p-1} \sqrt{\sum_{j=1}^n [x_j(t_{i+1}) - x_j(t_i)]^2}.$$

Les fonctions $x_j(t)$ étant lisses par morceaux, on a

$$x_j(t_{i+1}) - x_j(t_i) = \int_{t_i}^{t_{i+1}} x'_j(t) dt.$$

Mettons cette intégrale sous la forme $x'_j(t_i) \Delta t_i + \varepsilon_{ij} \Delta t_i$. Si t_i est un point de discontinuité de la fonction $x'_j(t)$, on sous-entend par $x'_j(t_i)$ la quantité $x'_j(t_i + 0)$. Il en résulte

$$|x_j(t_{i+1}) - x_j(t_i)| - |x'_j(t_i)| \Delta t_i \leq |\varepsilon_{ij}| \Delta t_i.$$

En vertu de l'inégalité 2.68 (6), on a

$$\left| \sqrt{\sum_{j=1}^n [x_j(t_{i+1}) - x_j(t_i)]^2} - \sqrt{\sum_{j=1}^n |x'_j(t_i)|^2 \Delta t_i} \right| \leq \sqrt{\sum_{j=1}^n |\varepsilon_{ij}|^2 \Delta t_i}.$$

Ensuite

$$\left| \sum_{i=0}^{p-1} \sqrt{\sum_{j=1}^n [x_j(t_{i+1}) - x_j(t_i)]^2} - \sum_{i=0}^{p-1} \sqrt{\sum_{j=1}^n |x'_j(t_i)|^2 \Delta t_i} \right| \leq \sum_{i=0}^{p-1} \sqrt{\sum_{j=1}^n |\varepsilon_{ij}|^2 \Delta t_i}. \quad (3)$$

La quantité ε_{ij} admet la majoration :

$$|\varepsilon_{ij}| = \left| \frac{1}{\Delta t_i} \int_{t_i}^{t_{i+1}} |x_j(t) - x_j(t_i)| dt \right| \leq \max_{t_i \leq t \leq t_{i+1}} |x_j(t) - x_j(t_i)|. \quad (4)$$

Partageons la somme dans le second membre de (3) en deux parties : la première comprendra les termes correspondant aux intervalles de continuité de toutes les $x_j(t)$, la deuxième les termes correspondant aux intervalles contenant les points de discontinuité d'au moins une des fonctions $x_j(t)$. Comme les fonctions $x_j(t)$ sont uniformément continues dans leurs intervalles de continuité, si (t_i, t_{i+1}) est un intervalle de continuité de la fonction $x_j(t)$, on a $\max |x_j(t) - x_j(t_i)| \leq \omega(d(\Pi))$, où $\omega(u)$ est infiniment petit avec u , $d(\Pi)$ étant, comme toujours, $\max \Delta t_i$. Si la fonction $x_j(t)$ subit une discontinuité dans l'intervalle Δt_i , alors le second membre de (4) admet la majoration

$$\max |x_j(t) - x_j(t_i)| \cdot \Delta t_i \leq M \cdot \Delta t_i,$$

où M est une constante. Définitivement, le second membre de (3) aura pour majorant

$$\omega(d(\Pi))(b-a) + M \sum' \Delta t_i,$$

où $\sum'$ se rapporte aux intervalles contenant les points de discontinuité d'au moins une des fonctions $x_j(t)$. Le nombre de ces points étant fixe, la somme $\sum'$ tend vers zéro avec $d(\Pi)$; il en est de même de $\omega(d(\Pi))$. D'autre part, la fonction

$$\sqrt{\sum_{j=1}^n |x_j(t)|^2}$$

est continue par morceaux elle aussi, et l'expression

$$\sum_{i=0}^{p-1} \sqrt{\sum_{j=1}^n |x_j(t_i)|^2} \Delta t_i$$

est sa somme intégrale qui tend vers

$$s = \int_a^b \sqrt{\sum_{j=1}^n |x_j(t)|^2} dt \quad (5)$$

lorsque $d(\Pi) \rightarrow 0$. Il résulte maintenant de (3) que L_p tend vers l'intégrale (5) lorsque $d(\Pi) \rightarrow 0$.

Définitivement, la longueur de l'arc de la courbe (2) déterminé par les valeurs $t = a$ et $t = b$ existe et s'exprime par l'intégrale (5),

cette longueur étant définie comme limite des lignes polygonales inscrites, et les fonctions $x_j(t)$ étant lisses par morceaux.

Pour $n = 2$, on a la formule ($x_1 = x$, $x_2 = y$):

$$s = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2} dt. \quad (6)$$

Si l'on choisit pour paramètre la coordonnée x , on retrouve la formule (1).

Exemples.

a. La longueur de l'arc de la chaînette $y = \operatorname{ch} x$, déterminé par les points $x = 0$ et $x = b$ (fig. 9.27) est égale à

$$s = \int_0^b \sqrt{1 + \operatorname{sh}^2 x} dx = \int_0^b \operatorname{ch} x dx = \operatorname{sh} b.$$

b. La longueur de l'arc de la circonférence $x = r \cos \theta$, $y = r \sin \theta$ intercepté par deux rayons $\theta = \theta_1$ et $\theta = \theta_2$, $0 \leq \theta_1 < \theta_2 \leq 2\pi$, est égale à

$$s = \int_{\theta_1}^{\theta_2} \sqrt{r^2 \sin^2 \theta + r^2 \cos^2 \theta} d\theta = r(\theta_2 - \theta_1)$$

d'après (6). En particulier, lorsque $\theta_1 = 0$, $\theta_2 = 2\pi$, on voit que la longueur de la circonférence de rayon r vaut $2\pi r$. Le nombre π reçoit donc une nouvelle interprétation: il représente le rapport de la longueur d'une circonférence à son diamètre.

c. La longueur de l'arc de l'ellipse

$$x = a \cos t, \quad y = b \sin t, \quad b > a$$

d'excentricité $e = \frac{c}{b} = \frac{\sqrt{b^2 - a^2}}{b}$, pour l'intervalle $0 \leq t \leq T$, est

$$\begin{aligned} s &= \int_0^T \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = \int_0^T \sqrt{a^2 \sin^2 t + b^2 (1 - \sin^2 t)} dt = \\ &= \int_0^T \sqrt{b^2 - (b^2 - a^2) \sin^2 t} dt = b \int_0^T \sqrt{1 - e^2 \sin^2 t} dt = b \cdot E(e, T), \end{aligned}$$

où $E(e, T)$ est la fonction elliptique de Legendre (9.45).

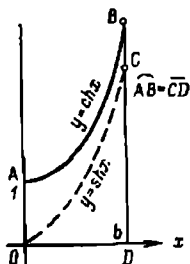


Fig. 9.27.

d. La longueur de la sinusoïde $y = \sin x$, pour $0 \leq x \leq \frac{\pi}{2}$, s'exprime elle aussi par une intégrale elliptique

$$\begin{aligned} s &= \int_0^{\pi/2} \sqrt{1 + \cos^2 x} \, dx = \int_0^{\pi/2} \sqrt{2 - \sin^2 x} \, dx = \\ &= \sqrt{2} \int_0^{\pi/2} \sqrt{1 - \frac{1}{2} \sin^2 x} \, dx = \sqrt{2} E\left(\frac{1}{\sqrt{2}}, \frac{\pi}{2}\right). \end{aligned}$$

Ce n'est d'ailleurs pas étonnant car l'ellipse que l'on obtient en coupant un cylindre circulaire droit par un plan incliné devient sinusoïde en développant ce cylindre dans le plan.

e. Contrairement à l'aire, la longueur d'une courbe ne se transforme en général pas d'une façon tant soit peu simple, lors d'une dilatation suivant un axe; même pour une ligne polygonale, la longueur de chaque segment se multiplie alors par son propre coefficient qui dépend de l'angle formé par ce segment et l'axe de dilatation.

Par contre, si la dilatation d'un même rapport k s'effectue suivant tous les axes à la fois (c'est-à-dire il s'agit d'une homothétie), alors la longueur de tout segment, de toute ligne polygonale, donc de toute courbe lisse par morceaux se multiplie par k .

f. Désignons par s_0 la longueur de la corde γ_0 joignant deux points $A = M_0$ et $B = M_p$ d'une courbe γ lisse par morceaux. Comme, en vertu de l'inégalité de triangle, toute ligne polygonale $M_0 M_1 \dots M_p$ a une longueur $s_p \geq s_0$, il en est de même de la longueur s de la courbe γ :

$$s \geq s_0.$$

Ainsi,

$$\frac{s}{s_0} \geq 1.$$

D'autre part, si la corde γ_0 se réduit au point M_0 , de sorte que le point B tend vers le point A , alors le rapport

$$\begin{aligned} \frac{s}{s_0} &= \frac{\int_a^b \sqrt{\sum_{j=1}^n [x'_j(t)]^2} \, dt}{\sqrt{\sum_{j=1}^n [x_j(b) - x_j(a)]^2}} = \\ &= \frac{(b-a) \sqrt{\sum_{j=1}^n [x'_j(\xi)]^2}}{(b-a) \sqrt{\sum_{j=1}^n [x'_j(\xi_j)]^2}} = \sqrt{\frac{\sum_{j=1}^n [x'_j(\xi)]^2}{\sum_{j=1}^n [x'_j(\xi_j)]^2}} \end{aligned}$$

a 1 pour limite si $\sum_{j=1}^n [x'_j(a)]^2 \neq 0$.

Nous aboutissons à un résultat important :

La longueur d'une courbe γ lisse par morceaux est un infiniment petit équivalent à la longueur de la corde qui sous-tend la courbe lorsque l'accroissement du paramètre tend vers 0.

g. Longueur d'arc comme fonction du paramètre et comme paramètre. Soit une courbe L donnée par ses équations paramétriques

$$x_i = x_i(\tau), \dots, x_n = x_n(\tau), \quad a \leq \tau \leq b,$$

où $x_j(\tau)$ sont des fonctions lisses par morceaux. La longueur de la courbe L comprise entre $\tau = a$ et $\tau = t$ est une fonction lisse par morceaux

$$s(t) = \int_a^t \sqrt{\sum_{j=1}^n [x_j'(\tau)]^2} d\tau$$

du paramètre t . Sa dérivée en t , à l'extérieur des points de discontinuité de $x_j'(t)$, a la forme

$$s'(t) = \sqrt{\sum_{j=1}^n [x_j'(t)]^2} \quad (7)$$

et est continue et non négative, de sorte que $s(t)$ ne décroît pas. Si $x_1'(t), \dots, x_n'(t)$ ne s'annulent simultanément en aucun point (pour un point de discontinuité, ce sont $x_j'(t+0)$ et $x_j'(t-0)$ qui ne s'annulent pas à la fois), alors, pour chaque intervalle de continuité de toutes les fonctions $x_j(t)$, on peut appliquer le théorème sur la fonction inverse: il existe la fonction inverse $t = t(s)$ continue, croissante et possédant une dérivée qui, d'après la formule 7.16 (2), est continue par morceaux de même que $s'(t)$. Ceci étant, toutes les fonctions $x_j(t)$ peuvent être mises sous forme de fonctions de s , continues et possédant des dérivées continues par morceaux. En particulier, si t désigne la longueur d'arc, alors il résulte de (7) que

$$\sum_{j=1}^n [x_j'(s)]^2 = 1. \quad (8)$$

Un point où toutes les fonctions $x_j'(t)$ s'annulent s'appelle *point singulier*; nous voyons que la représentation paramétrique, avec la longueur d'arc pour paramètre, n'est possible que sur une partie de la courbe sans points singuliers.

9.64. Aire d'une surface de révolution. Soit une courbe $L = \{y = f(x) \geq 0, x \in [a, b]\}$. En tournant autour de l'axe des x , elle engendre une surface P dont nous voulons déterminer l'aire (fig. 9.28).

Supposons d'abord que la fonction $f(x)$ soit du premier degré, $f(x) = kx + b_1$; la surface P est alors appelée *cône* (fig. 9.29); la ligne $y = kx + b_1$ est la *génératrice* du cône, le point A d'intersection de la génératrice et de l'axe son *sommet*, la circonférence γ décrite par le point B lors de sa révolution sa *circonférence directrice*. L'aire du cône est définie de la façon suivante. Choisissons, sur la circonférence γ , les points $B_1, \dots, B_m$ et joignons-les successivement l'un avec l'autre et chacun avec le sommet A ; on obtient une ligne polygonale fermée $\{B_1 \dots B_m B_1\} = \lambda$ et un système $B_1 A B_2, \dots, B_m A B_1$

de triangles isocèles formant une *pyramide inscrite* Π (fig. 9.30). L'aire de la pyramide est définie comme somme $S(\Pi)$ des aires des triangles qui la forment. Posons $\overline{B_k B_{k+1}} = 2\delta_k$ ($k =$

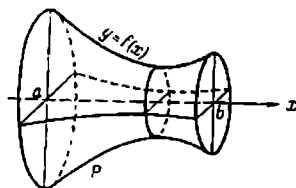


Fig. 9.28.

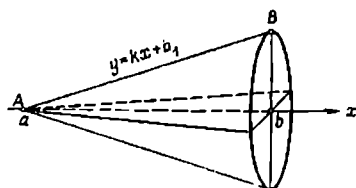


Fig. 9.29.

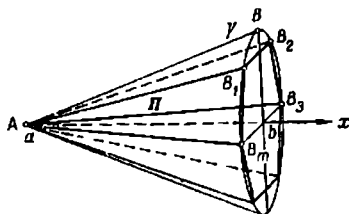


Fig. 9.30.

$= 1, \dots, m)$ et $d(\Pi) = \max_k \delta_k$. Considérons une suite quelconque de pyramides inscrites $\Pi_1, \dots, \Pi_n, \dots$ avec $d(\Pi_n) \rightarrow 0$. Il se trouve que les nombres $S(\Pi_n)$ tendent vers une limite qui ne dépend pas du choix de ces pyramides; cette limite est, par définition, l'aire du cône P .

Pour démontrer l'existence de la limite de la suite $S(\Pi_n)$, explicitons les nombres $S(\Pi)$. Si l est la longueur de la génératrice du cône (du sommet au point B), la hauteur du triangle $B_k A B_{k+1}$ vaut

$$\sqrt{l^2 - \delta_k^2} = l \sqrt{1 - \left(\frac{\delta_k}{l}\right)^2} = l \left(1 - \frac{\delta_k^2 E}{2l^2}\right) = l - \frac{\delta_k^2 E}{2l},$$

où E a 1 pour limite, lorsque $\delta_k \rightarrow 0$ (5.59.f). L'aire du triangle $B_k A B_{k+1}$ est égale à $\delta_k \sqrt{l^2 - \delta_k^2} = \delta_k l - \frac{\delta_k^3 E}{2l}$ et celle de la pyramide inscrite à

$$S(\Pi) = \sum_k \left(\delta_k l - \frac{\delta_k^3 E}{2l} \right) = \frac{1}{2} l \beta(\Pi) - \frac{1}{2l} \sum_k \delta_k^3 E, \quad (1)$$

où $\beta(\Pi)$ est la longueur de la ligne polygonale λ . Pour $d(\Pi) \rightarrow 0$, la quantité $\beta(\Pi)$ tend vers la longueur de la circonférence γ (9.63.b) qui vaut $2\pi R$, de sorte que le premier terme à droite dans (1) a $\pi l R$ pour limite. Si chaque $\delta_k < e$, alors le deuxième terme à droite dans (1) est majoré par $\frac{1}{2l} \cdot 2e^3 \beta(\Pi)$ (pour δ_k assez petits); il en résulte que ce terme tend vers 0 pour $d(\Pi) \rightarrow 0$. Ainsi, la suite numérique $S(\Pi_n)$ tend vers $\pi l R$ qui est, on l'a vu, l'aire du cône.

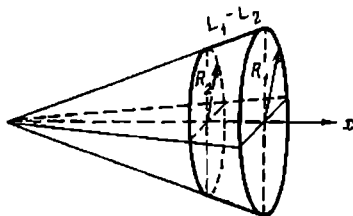


Fig. 9.31.

L'aire d'un cône tronqué (fig. 9.31) engendrée par la révolution

autour de l'axe des x d'un segment de la droite $y = kx + b$ est définie comme différence des aires de deux cônes aux génératrices L_1 et L_2 et aux rayons de base R_1 et R_2 respectivement. On aboutit à la quantité

$$\pi (L_1 R_1 - L_2 R_2) = \pi (L_1 - L_2) (R_1 + R_2)$$

puisque $L_1 R_2 = L_2 R_1$, ou bien, ce qui revient au même, $\frac{L_1}{L_2} = \frac{R_1}{R_2}$. Ceci résulte immédiatement de l'homothétie. Le nombre $L_1 - L_2$ est, évidemment, la longueur du segment de la génératrice du cône.

Passons au cas général et, à la place de $y = y(x)$, considérons une courbe sous forme paramétrique

$$L = \{x = x(t), y = y(t) \geq 0, \alpha \leq t \leq \beta, x(\alpha) = a, x(\beta) = b\}.$$

Définissons l'aire de la surface engendrée par la révolution autour de l'axe des x de cette courbe comme limite des aires des surfaces engendrées par la révolution autour de l'axe des x des lignes polygonales inscrites dans la courbe L . Soient une partition $\Pi = \{\alpha = t_0 \leq t_1 \leq \dots \leq t_n = \beta\}$ de l'intervalle $[\alpha, \beta]$ et L_n la ligne polygonale à sommets $(x(t_0), y(t_0)), \dots, (x(t_n), y(t_n))$; la longueur du k -ième côté de cette ligne est $\Delta l_k = \sqrt{\Delta x_k^2 + \Delta y_k^2}$, où $\Delta x_k = x(t_{k+1}) - x(t_k)$, $\Delta y_k = y(t_{k+1}) - y(t_k)$. En tournant autour de l'axe des x , le k -ième côté engendre le tronc de cône dont l'aire

vaut $\pi (y_k + y_{k+1}) \Delta t_k$; en sommant suivant tous les côtés de la ligne on obtient l'expression de l'aire engendrée par la révolution de la ligne polygonale:

$$P_n = \pi \sum_{k=0}^{n-1} (y_k + y_{k+1}) \Delta t_k.$$

Supposons maintenant que les fonctions $x(t)$ et $y(t)$ aient leurs dérivées continues $x'(t)$ et $y'(t)$. On peut alors transformer les expressions Δx_k et Δy_k à l'aide de la formule de Lagrange 7.44 en aboutissant à

$$P_n = \pi \sum_{k=0}^{n-1} [y(t_k) + y(t_{k+1})] \sqrt{[x'(\xi_k)]^2 + [y'(\eta_k)]^2} \Delta t_k,$$

les points ξ_k et η_k se trouvant dans l'intervalle (t_k, t_{k+1}) . A côté de l'expression P_n , considérons la quantité

$$P_n^* = 2\pi \sum_{k=0}^{n-1} y(t_k) \sqrt{[x'(t_k)]^2 + [y'(t_k)]^2} \Delta t_k.$$

P_n^* est une somme intégrale pour la fonction continue $2\pi y(t) \sqrt{[x'(t)]^2 + [y'(t)]^2}$ et tend, lorsque $d(\Pi) \rightarrow 0$, vers le nombre

$$S = 2\pi \int_a^b y(t) \sqrt{[x'(t)]^2 + [y'(t)]^2} dt. \quad (2)$$

Montrons que P_n a la même limite suivant la même direction. Il suffit de prouver que

$$\lim_{d(\Pi) \rightarrow 0} [P_n^* - P_n] = 0.$$

En effet, pour un $\varepsilon > 0$ donné, on peut trouver un $\delta > 0$ tel que $|t' - t| < \delta$ implique les inégalités $|x'(t') - x'(t)| < \varepsilon$, $|y'(t') - y'(t)| < \varepsilon$, $|y(t') - y(t)| < \varepsilon$. Si $d(\Pi) < \delta$, alors, en particulier,

$$\begin{aligned} y(t_{k+1}) &= y(t_k) + \theta_1 \varepsilon, & |\theta_1| < 1, \\ x'(\xi_k) &= x'(t_k) + \theta_2 \varepsilon, & |\theta_2| < 1, \\ y'(\eta_k) &= y'(t_k) + \theta_3 \varepsilon, & |\theta_3| < 1 \end{aligned}$$

de façon que

$$\begin{aligned} & [y(t_k) + y(t_{k+1})] \sqrt{[x'(\xi_k)]^2 + [y'(\eta_k)]^2} = \\ &= (2y(t_k) + \theta_1 \varepsilon) \sqrt{[x'(t_k) + \theta_2 \varepsilon]^2 + [y'(t_k) + \theta_3 \varepsilon]^2} = \\ &= (2y(t_k) + \theta_1 \varepsilon) \sqrt{[x'(t_k)]^2 + [y'(t_k)]^2 + \varepsilon \theta_4}, \end{aligned}$$

où $|\theta_4| \leq 2 \max |x'(t)| + |y'(t)| + \varepsilon$. Quels que soient a , $b > 0$, on a les inégalités

$$\sqrt{a} - \sqrt{b} \leq \sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$$

(la première est évidente, la seconde est l'inégalité de triangle dans R_2 pour les vecteurs $(\sqrt{a}, 0)$ et $(0, \sqrt{b})$). Il en résulte que

$$\sqrt{a+b} = \sqrt{a} + \theta_5 \sqrt{b}, \quad |\theta_5| \leq 1.$$

Dans le présent cas, on peut noter

$$\sqrt{|x'(t_k)|^2 + |y'(t_k)|^2 + \varepsilon \theta_4} = \sqrt{|x'(t_k)|^2 + |y'(t_k)|^2} + \theta_5 \sqrt{\varepsilon \theta_4},$$

où $|\theta_5| \leq 1$. Donc

$$\begin{aligned} P_n - P_n^* &= -2\pi \sum_{k=0}^{n-1} y(t_k) \sqrt{|x'(t_k)|^2 + |y'(t_k)|^2} + \\ &+ \pi \sum_{k=0}^{n-1} (2y(t_k) + \theta_4 \varepsilon) [\sqrt{|x'(t_k)|^2 + |y'(t_k)|^2} + \theta_5 \sqrt{\varepsilon \theta_4}] = \\ &= \sum_{k=0}^{n-1} (\pi \theta_4 \varepsilon \sqrt{|x'(t_k)|^2 + |y'(t_k)|^2} + 2\pi y(t_k) \theta_5 \sqrt{\varepsilon \theta_4} + \\ &\quad + \pi \theta_4 \theta_5 \varepsilon \sqrt{\varepsilon \theta_4}). \end{aligned}$$

Il est évident que cette quantité tend vers zéro avec ε , ce qu'il fallait démontrer.

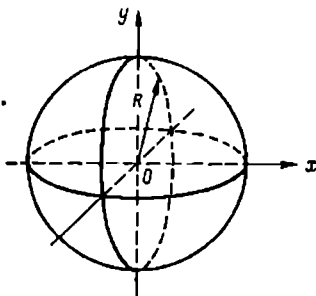


Fig. 9.32.

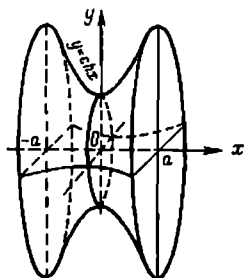


Fig. 9.33.

Définitivement, la formule (2) fournit l'expression de l'aire cherchée dans l'hypothèse d'existence et de continuité de $y'(t)$ et $x'(t)$. Pour $t \equiv x$, la formule se met sous la forme

$$S = 2\pi \int_a^b y(x) \sqrt{1 + [y'(x)]^2} dx.$$

E x e m p l e s.

a. Sphère de rayon R . La sphère en question est la surface engendrée par la demi-circonférence $x = R \cos \theta$, $y = R \sin \theta$, $0 \leq \theta \leq \pi$, tournant autour de l'axe des x (fig. 9.32). Donc, d'après la formule (2), on a

$$S = 2\pi \int_0^{\pi} R \sin \theta \sqrt{R^2 \sin^2 \theta + R^2 \cos^2 \theta} d\theta = 2\pi R^2 \int_0^{\pi} \sin \theta d\theta = 4\pi R^2.$$

b. La caténoïde est la surface engendrée par la révolution autour de l'axe des x d'une chaînette (fig. 9.33). Nous avons $y = \operatorname{ch} x$, $y' = \operatorname{sh} x$, $\sqrt{1+y'^2} = \operatorname{ch} x$,

$$S = 2\pi \int_{-a}^{+a} \operatorname{ch}^2 x dx = 2\pi \int_{-a}^{+a} \frac{\operatorname{ch} 2x + 1}{2} dx = \pi \left[\frac{\operatorname{sh} 2x}{2} + x \right] \Big|_{-a}^{+a} = \pi (\operatorname{sh} 2a + 2a).$$

9.65. Schéma général d'application de l'intégrale. Soit une quantité v à calculer. Supposons qu'elle représente l'une des valeurs d'une fonction $v(x)$ de la variable x parcourant un intervalle $[a, b]$, par exemple $v = v(b)$. Si l'on peut dégager la partie linéaire principale

$$dv = f(x) dx \quad (1)$$

de l'accroissement Δv de la fonction $v(x)$, pour le passage de x à $x + dx$, alors le problème se réduit en principe au calcul de l'intégrale

$$v(b) = v(a) + \int_a^b f(x) dx \quad (2)$$

pour $v(a)$ connue.

En effet, le coefficient $f(x)$ dans la partie linéaire principale de l'accroissement est la dérivée de la fonction $v(x)$, et le problème consiste à retrouver une primitive, ce qu'on fait à l'aide de l'intégrale (2).

Dans certains cas où il s'avère difficile de construire l'égalité (1), mais que l'on peut établir une inégalité de la forme

$$dv \leq f(x) dx,$$

l'intégration de celle-ci fournit une estimation de la quantité $v(b)$:

$$v(b) \leq v(a) + \int_a^b f(x) dx.$$

Notons qu'en écrivant l'expression de dv (contrairement à celle de Δv), on peut négliger les infiniment petits d'ordre supérieur par rapport à dx . Cette considération simplifie souvent le problème.

Voici quelques exemples simples.

u. Quel travail faut-il accomplir pour faire monter une masse de 1 kg se trouvant sur la surface de la Terre hors les limites de la gravitation terrestre?

Désignons par $W(h)$ le travail fourni pour remonter 1 kg à une altitude h . La force d'attraction de la Terre $F = F(h)$ sur la masse de 1 kg à l'altitude h m est égale, selon la loi de Newton, à $C \cdot \frac{1}{(h+R)^2}$, où R est le rayon de la Terre.

On détermine la constante C à partir de la condition qu'au niveau de la mer la force attractive subie par cette masse est égale à 1 kg de poids ($= 9,8 \text{ kg} \cdot \text{m/s}^2$).

Donc $C = R^2$, $F(h) = \frac{R^2}{(h+R)^2} \text{ kg}$. Supposons que la masse s'élève à l'altitude $(h + \Delta h) \text{ m}$. En admettant que la force $F(h)$ soit constante sur cet intervalle, on obtient pour le travail

$$dW = F dh = \frac{R^2}{(h+R)^2} dh \text{ kg} \cdot \text{m}.$$

En réalité, la force F n'est pas constante sur l'intervalle $[h, h + \Delta h]$, mais si l'on en tenait compte, ceci n'ajouterait à l'expression de dW que des infiniment petits d'ordre supérieur ($\sim \Delta F \cdot \Delta h$). Or, nous n'avons besoin que des infiniment petits du premier ordre par rapport à dh .

A présent, l'application de la formule (1) fournit la valeur du travail qu'il faut accomplir pour faire monter 1 kg à une altitude H :

$$W = \int_{h=0}^H \frac{R^2}{(h+R)^2} dh = \frac{R^2}{h+R} \Big|_0^H = R - \frac{R^2}{H+R} < R = 6,4 \cdot 10^6 \text{ kg} \cdot \text{m}.$$

Pour H très grand, on peut négliger la quantité $\frac{R^2}{H+R}$ (c'est dans ce cas que l'on dit « hors les limites de la gravitation terrestre », parce que l'ascension ne demande pratiquement plus de travail à accomplir pour vaincre la force attractive de la Terre). La réponse est donc $W = 6,4 \cdot 10^6 \text{ kg} \cdot \text{m}$.

b. Un entonnoir de section supérieure S , de section inférieure s et de hauteur H est rempli d'eau (fig. 9.34). On sait que la vitesse d'écoulement sous le poids d'une colonne d'eau de hauteur h est

$$v(h) = \sqrt{2gh} \quad (g = 9,81 \text{ m/s}^2).$$

En combien de temps l'entonnoir va-t-il se vider?

Désignons par $S(h)$ la section de l'entonnoir à la hauteur $h = h(t)$ et par $v = v(t)$ le volume d'eau dans l'entonnoir à l'instant t . L'eau qui s'écoule par l'orifice inférieur durant l'intervalle de temps $[t, t + \Delta t]$ est $v \cdot s \cdot dt = \sqrt{2gh} \cdot s \cdot dt \text{ cm}^3$ (la vitesse d'écoulement est supposée constante durant ce temps, l'erreur qui s'ensuit est un infiniment petit d'ordre supérieur que l'on néglige). La même quantité dv s'obtient en admettant que le volume d'eau qui s'écoule soit égal au volume du cylindre de hauteur dh et de base $S(h)$ (en réalité, ce volume est celui d'un tronc de cône, mais la différence est infiniment petite d'ordre supérieur par rapport à dh). On aboutit à l'égalité

$$S(h) dh = \sqrt{2gh} \cdot s \cdot dt,$$

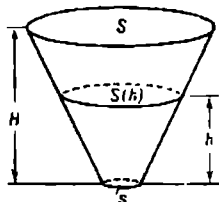


Fig. 9.34.

d'où

$$s \cdot dt = \frac{S(h)}{\sqrt{2gh}} dh = \frac{h^2}{H^2} S \frac{dh}{\sqrt{2gh}} = \frac{S}{H^2} \sqrt{\frac{h^3}{2g}} dh.$$

En intégrant en t de 0 à T et en tenant compte de 9.54 on a

$$s \int_0^T dt = sT = \frac{S}{H^2 \sqrt{2g}} \int_0^H \sqrt{h^3} dh = \frac{S}{H^2 \sqrt{2g}} \frac{h^{5/2}}{5/2} \Big|_0^H = \frac{2}{5} \sqrt{\frac{H}{2g}} S,$$

d'où

$$T = \frac{2}{5} \frac{S}{s} \sqrt{\frac{H}{2g}}.$$

c. Estimation de l'aire du ρ -voisinage d'une courbe plane. Le lieu géométrique $V_\rho(L)$ des points de tous les cercles de rayon ρ dont les centres appartiennent à une

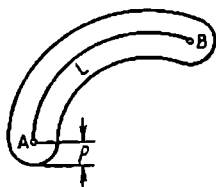


Fig. 9.35.

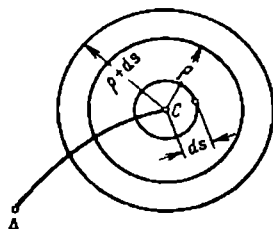


Fig. 9.36.

courbe L s'appelle ρ -voisinage de la courbe L (fig. 9.35). Supposons que la courbe L soit lisse par morceaux; on peut démontrer que l'ensemble $V_\rho(L)$ possède alors une aire. Nous allons montrer que l'aire $S_\rho(L)$ de l'ensemble $V_\rho(L)$ admet l'estimation

$$S_\rho(L) \leq 2\pi\rho s(L) + \pi\rho^2, \quad (2)$$

où $s(L)$ est la longueur de la courbe L .

Choisissons pour paramètre, sur la courbe L , la longueur d'arc s comptée à partir du point A ($s = 0$) jusqu'au point B ($s = s_0$) (9.63.g). Désignons par $S_\rho(L_\sigma)$ l'aire du ρ -voisinage de la partie de la courbe L correspondant à l'intervalle $0 \leq s \leq \sigma$ de variation du paramètre s . Soit Q_ρ le cercle de rayon ρ centré en C ($s = \sigma$). Les positions possibles des points $C' \in L$ qui correspondent aux valeurs du paramètre s dans l'intervalle $\sigma \leq s \leq \sigma + ds$ appartiennent au cercle de rayon ds et de centre C (fig. 9.36). Les positions possibles des points des cercles de rayon ρ et de centre C' appartiennent au cercle $Q_{\rho+ds}$ de rayon $\rho + ds$ centré en C . Par conséquent,

l'accroissement total de $S_p(L_s)$, lorsque le paramètre s varie de σ à $\sigma + ds$, ne dépasse pas la différence des aires des cercles $Q_{\sigma+ds}$ et Q_σ , c'est-à-dire la quantité

$$\pi(\rho + ds)^2 - \pi\rho^2 = 2\pi\rho ds + \pi(ds)^2.$$

Pour la partie linéaire principale de cet accroissement, on obtient l'estimation

$$dS_p(L_s) \leq 2\pi\rho ds.$$

En intégrant en s de 0 à s_0 , on trouve

$$S_p(L) - S_p(L_0) \leq 2\pi\rho s(L). \quad (3)$$

Puisque $S_p(L_0) = \pi\rho^2$, (2) résulte de (3).

d. Volume d'un corps à bases parallèles. Il résulte de la théorie générale du calcul des volumes que le volume

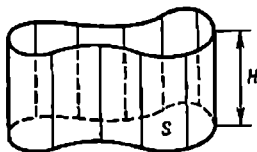


Fig. 9.37

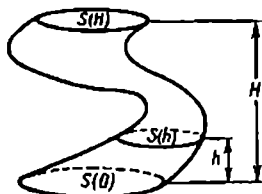


Fig. 9.38.

du cylindre droit de base jordanienne S et de hauteur H est égal à SH (fig. 9.37). Soit un corps quelconque dont la section à la hauteur h est jordanienne et dont l'aire est $S(h)$ (fig. 9.38). Supposons que la forme de la section V_h varie continûment avec h ; ceci veut dire que, pour tout $\epsilon > 0$, il existe un $\delta > 0$ tel que si $|h - h'| < \delta$, la projection de la frontière de $V_{h'}$ sur le plan de la base appartient au ϵ -voisinage de la projection de la frontière de V_h sur le même plan (c). La différence des aires des sections V_h et $V_{h'}$ ne dépasse donc pas l'aire du voisinage mentionné et admet pour majoration, d'après c, le nombre $C\epsilon L$, où L est un majorant des longueurs des contours de toutes les sections horizontales.

Désignons par $v(h)$ le volume du corps limité en haut par la section V_h . En passant de h à $h + dh$, la partie linéaire principale de l'accroissement de $v(h)$ se met sous la forme $S(h) dh$ comme si le corps était un cylindre dans l'intervalle $[h, h + dh]$; l'accroissement véritable diffère de cette quantité par un infiniment petit d'ordre supérieur par rapport à dh . Donc,

$$dv = S(h) dh.$$

En intégrant on obtient

$$v = \int_0^H S(h) dh. \quad (4)$$

e. Volume d'un cône K de hauteur H et de base supérieure S (fig. 9.39). Pour la raison de similitude, l'aire de la section à une distance h du sommet est $S(h) = S \left(\frac{h}{H} \right)^2$, d'où

$$v(K) = \int_0^H S \frac{h^2}{H^2} dh = \frac{S}{H^2} \frac{1}{3} H^3 = \frac{1}{3} SH.$$

f. Volume d'une boule de rayon R . L'aire de la

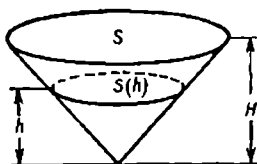


Fig. 9.39.

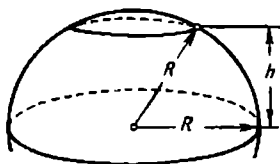


Fig. 9.40.

section distante de h du centre est $\pi (R^2 - h^2)$ (fig. 9.40), d'où

$$\begin{aligned} v &= 2 \int_0^R \pi (R^2 - h^2) dh = 2\pi \left[R^2 h - \frac{h^3}{3} \right] \Big|_0^R = \\ &= 2\pi \left(R^3 - \frac{R^3}{3} \right) = \frac{4}{3} \pi R^3. \end{aligned}$$

g. Pour un corps de révolution, la formule (4) peut être mise sous la forme

$$v = \pi \int_0^H r^2(h) dh, \quad (5)$$

$r(h)$ étant le rayon de la section à une hauteur h .

h. Volume d'un corps commun à une boule de rayon R et à un parabolôïde dont le sommet coïncide avec le centre de la boule (fig. 9.41). L'équation du parabolôïde est $2az = x^2 + y^2$; il en résulte que l'aire de la section du parabolôïde par un plan à une distance h (du centre de la boule) vaut $\pi \cdot 2ah$. L'aire de la section de la boule à la hauteur h est $\pi (R^2 - h^2)$. Les surfaces de la boule et du parabolôïde se coupent à la hauteur z_0 déterminée de l'équation

$2az_0 = R^2 - z_0^2$; donc, $z_0 = -a + \sqrt{a^2 + R^2}$ et

$$v = \int_0^{z_0} \pi \cdot 2ah \, dh + \int_{z_0}^R \pi (R^2 - h^2) \, dh.$$

Le lecteur calculera sans peine les intégrales obtenues.

i. A la suite d'une dilatation de rapport k suivant un axe, le (volume d'un corps est multiplié par k démonstration analogue à celle de p.61.d-e pour le cas plan). Il en résulte, par exemple, que le volume d'un ellip-

soïde de demi-axes a, b, c est $\frac{4}{3} \pi abc$,

parce que la boule de rayon 1 a $\frac{4}{3} \pi$ pour

volume, et l'ellipsoïde s'en déduit par une dilatation de rapport a suivant l'axe des x , une dilatation de rapport b suivant l'axe des y et une dilatation de rapport c suivant l'axe des z .

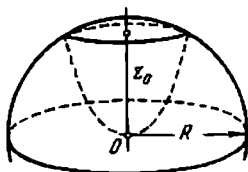


Fig. 9.41.

§ 9.7. Intégration et dérivation d'une suite de fonctions

9.71. Soit une suite $f_1(x), f_2(x), \dots, f_n(x), \dots$ de fonctions intégrables sur un intervalle $a \leq x \leq b$ qui est partout convergente vers une fonction $f(x)$ sur $[a, b]$.

(a) Est-ce que la fonction $f(x)$ est intégrable elle aussi ?

(b) Si oui, est-ce que la relation

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx \quad (1)$$

est valable ?

Il s'avère que, dans le cas général, les réponses aux deux questions sont négatives.

(a) Supposons que $f_n(x)$ vaille 1 si $x \in [0, 1]$ est de la forme $\frac{p}{q}$, où $q \leq n$, et 0 dans le cas contraire. Ainsi, la fonction $f_n(x)$ n'est non nulle qu'en un nombre fini de points, donc (9.16.c) elle est intégrable (son intégrale est nulle). La suite $f_n(x)$ possède en tout point de l'intervalle $[0, 1]$, pour $n \rightarrow \infty$, la limite $f(x)$ égale à l'unité en chaque point rationnel et nulle en chaque point irrationnel. La fonction $f(x)$ n'est pas intégrable (9.17).

(b) Supposons que $f_n(x)$ vaille $n \sin nx$ pour $0 \leq x \leq \frac{\pi}{n}$ et 0 pour $\frac{\pi}{n} \leq x \leq \pi$. La suite $f_n(x)$ tend vers zéro en tout point

$x \in [0, \pi]$. La fonction limite est, évidemment, intégrable, et son intégrale sur l'intervalle $[0, \pi]$ est nulle. Pourtant,

$$\int_0^{\pi} f_n(x) dx = \int_0^{\pi/n} n \sin nx dx = 2,$$

et l'égalité (1) est fausse.

9.72. On peut s'attendre à ce que les réponses à 9.71 (a), (b) soient positives si la convergence de la suite $f_n(x)$ satisfait à quelques conditions supplémentaires. La convergence uniforme de la suite $f_n(x)$ vers sa limite $f(x)$ que nous avons définie dans 5.93 représente une telle condition supplémentaire. Rappelons la définition correspondante adaptée à notre cas: la suite $f_1(x), f_2(x), \dots$ converge uniformément sur $[a, b]$ vers la fonction $f(x)$ si, pour tout $\varepsilon > 0$, il existe un N tel que, quels que soient $x \in [a, b]$ et $n \geq N$, on a $|f(x) - f_n(x)| < \varepsilon$.

Théorème. Si une suite de fonctions intégrables $f_1(x), f_2(x), \dots$ converge uniformément sur $[a, b]$ vers une fonction $f(x)$, alors $f(x)$ est intégrable, et la relation

$$\lim_{n \rightarrow \infty} \int_a^x f_n(\xi) d\xi = \int_a^x f(\xi) d\xi$$

a lieu uniformément sur $[a, b]$. En particulier,

$$\lim_{n \rightarrow \infty} \int_a^b f_n(\xi) d\xi = \int_a^b f(\xi) d\xi.$$

Démonstration. Quelles que soient les partitions Π et Π' de l'intervalle $[a, b]$ et pour tout $n = 1, 2, \dots$, on a:

$$S_{\Pi}(f) = S_{\Pi}(f_n) + S_{\Pi}(f - f_n), \quad (1)$$

$$S_{\Pi'}(f) = S_{\Pi'}(f_n) + S_{\Pi'}(f - f_n). \quad (2)$$

Pour un $\varepsilon > 0$ donné, choisissons un numéro n de façon à avoir

$$|f(x) - f_n(x)| < \frac{\varepsilon}{4(b-a)}$$

partout sur $[a, b]$. Alors, nous avons évidemment

$$|S_{\Pi}(f - f_n)| \leq \frac{\varepsilon}{4}, \quad |S_{\Pi'}(f - f_n)| \leq \frac{\varepsilon}{4}.$$

La fonction $f_n(x)$ étant intégrable sur $[a, b]$, donc aussi sur $[a, x]$ (9.15.k), il existe un $\delta > 0$ tel que, quelles que soient deux partitions Π et Π' avec $d(\Pi) < \delta$, $d(\Pi') < \delta$, on a

$$|S_{\Pi}(f_n) - S_{\Pi'}(f_n)| < \frac{\varepsilon}{2}.$$

Il résulte alors de (1) et (2) :

$$\begin{aligned} |S_{\Pi}(f) - S_{\Pi'}(f)| &\leq |S_{\Pi}(f_n) - S_{\Pi'}(f_n)| + \\ &+ |S_{\Pi}(f - f_n)| + |S_{\Pi'}(f - f_n)| \leq \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4(b-a)}(b-a) + \frac{\varepsilon}{4(b-a)}(b-a) = \varepsilon. \end{aligned}$$

Cela signifie que les sommes intégrales de la fonction $f(x)$ satisfont au critère de Cauchy ; vu le théorème 4.19 appliqué à la direction $d(\Pi) \rightarrow 0$, il en résulte l'existence de l'intégrale.

Pour démontrer la seconde partie du théorème, notons que

$$\begin{aligned} \left| \int_a^x f(\xi) d\xi - \int_a^x f_n(\xi) d\xi \right| &= \left| \int_a^x |f(\xi) - f_n(\xi)| d\xi \right| \leq \\ &\leq \int_0^x |f(\xi) - f_n(\xi)| d\xi \leq \varepsilon(x-a) \leq \varepsilon(b-a) \end{aligned} \quad (3)$$

quel que soit le numéro n pourvu qu'il satisfasse à la relation

$$|f(x) - f_n(x)| \leq \varepsilon \quad (4)$$

pour tous les x . Comme la suite $f_n(x)$ converge uniformément vers $f(x)$, l'inégalité (4) a lieu pour tous les n à partir d'un certain N , ε étant fixe ; il en découle le résultat demandé.

9.73. Le théorème 9.72 démontré pour une suite de fonctions s'étend immédiatement au cas d'une série de fonctions :

T h é o r è m e. *Si une série de fonctions intégrables*

$$\varphi_1(x) + \varphi_2(x) + \dots + \varphi_n(x) + \dots \quad (1)$$

converge uniformément sur $[a, b]$ vers sa somme $S(x)$, alors $S(x)$ est intégrable, et la relation

$$\lim_{n \rightarrow \infty} \sum_{h=1}^n \int_a^x \varphi_h(\xi) d\xi = \int_a^x S(\xi) d\xi \quad (2)$$

a lieu iniformément sur $[a, b]$. En particulier,

$$\sum_{h=1}^{\infty} \int_a^b \varphi_h(\xi) d\xi = \int_a^b S(\xi) d\xi. \quad (3)$$

La démonstration résulte immédiatement de la convergence uniforme des sommes partielles de la série (1) vers $S(x)$ et du théorème 9.72.

9.74. Exemples. Les séries

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots, \quad (1)$$

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots \quad (2)$$

convergent uniformément sur tout intervalle $[-b, b]$, $0 < b < 1$. En les intégrant terme à terme de 0 à t on obtient encore deux séries

$$\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \dots, \quad (3)$$

$$\operatorname{arctg} t = t - \frac{t^3}{3} + \frac{t^5}{5} - \dots \quad (4)$$

convergentes uniformément sur $[-b, b]$, $0 < b < 1$. Etant donné que les séries dans les seconds membres de (3), (4) convergent aussi pour $t = 1$ et que les fonctions dans les premiers membres sont continues pour $0 \leq t \leq 1$, on a, compte tenu de 6.67 et 5.68 :

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \dots, \quad (5)$$

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \dots \quad (6)$$

9.75. On peut se servir du théorème 9.73 pour développer plusieurs fonctions non élémentaires en séries de puissances. Par exemple, l'intégration du développement

$$\frac{\sin t}{t} = 1 - \frac{t^2}{3!} + \frac{t^4}{5!} - \frac{t^6}{7!} + \dots$$

convergeant uniformément sur tout intervalle $[-T, T]$ fournit le développement du sinus intégral :

$$\text{Si } t = \int_0^t \frac{\sin \xi}{\xi} d\xi = t - \frac{t^3}{3 \cdot 3!} + \frac{t^5}{5 \cdot 5!} - \frac{t^7}{7 \cdot 7!} + \dots$$

convergeant uniformément lui aussi sur tout intervalle $[-T, T]$.

9.76. Formulons les questions analogues à 9.71 (a), (b) relativement à la dérivation. Soit une suite de fonctions dérivables sur un intervalle $a \leq x \leq b$:

$$f_1(x), f_2(x), \dots, f_n(x), \dots$$

cette suite étant partout convergente sur $[a, b]$ vers une fonction $f(x)$.

(a) Est-ce que la fonction $f(x)$ est dérivable?

(b) Si oui, est-ce que la relation

$$f'(x) = \lim f'_n(x) \quad (1)$$

a lieu?

La réponse à deux questions est toujours négative dans le cas général, même si la convergence de la suite $f_n(x)$ est uniforme :

(a) La suite de fonctions dérivables sur $[-1, +1]$:

$$f_n(x) = |x|^{1 + \frac{1}{n}}$$

converge uniformément sur $[-1, +1]$ vers la fonction $f(x) = |x|$ qui n'est pas dérivable pour $x = 0$.

(b) La suite

$$f_n(x) = \frac{1}{n} \sin nx$$

converge uniformément sur $[0, \pi]$ vers 0. La fonction limite est, évidemment, dérivable. Mais la relation (1) n'a pas lieu : la suite $f'_n(x) = \cos nx$ ne converge en aucun point sauf $x = 0$.

9.77. La situation change si l'on suppose la *convergence uniforme des dérivées*. Il s'avère alors que les conditions à imposer à la suite des fonctions $f_n(x)$ sont beaucoup moins gênantes :

T h é o r è m e. *Si une suite $x_n(t)$ de fonctions lisses par morceaux (9.39) converge au moins en un point t_0 d'un intervalle $[a, b]$ et la suite $x'_n(t)$ converge uniformément sur $[a, b]$ vers une fonction continue par morceaux $g(t)$, alors la suite $x_n(t)$ converge uniformément sur $[a, b]$ vers une fonction lisse par morceaux $x(t)$ qui a une dérivée $x'(t) = \lim_{n \rightarrow \infty} x'_n(t) = g(t)$ en tout point de continuité de la fonction $g(t)$.*

D é m o n s t r a t i o n. Pour une fonction $x_n(t)$, nous avons d'après 9.32 :

$$x_n(t) - x_n(t_0) = \int_{t_0}^t x'_n(\xi) d\xi. \quad (1)$$

Les fonctions $x_n(t) - x_n(t_0)$ convergent uniformément sur l'intervalle $[a, b]$ d'après le théorème 9.72 ; or les nombres $x_n(t_0)$ tendent vers une limite, donc la suite $x_n(t)$ converge uniformément. Désignons sa limite par $x(t)$. En passant à la limite dans l'égalité (1) et en se servant encore une fois du théorème 9.72 on obtient

$$x(t) - x(t_0) = \int_{t_0}^t g(\xi) d\xi.$$

La fonction $g(t)$ étant continue par morceaux, la fonction $x(t)$ est dérivable partout sauf aux points de discontinuité de $g(t)$ (9.31) ; de plus, $x'(t) = g(t) = \lim_{n \rightarrow \infty} x'_n(t)$ en tout point de continuité de $g(t)$, ce qu'il fallait démontrer.

9.78. Énonçons le théorème analogue pour une série.

T h é o r è m e. Si une série $u_1(x) + u_2(x) + \dots$ de fonctions lisses par morceaux converge au moins en un point d'un intervalle $[a, b]$ et si la série des dérivées $u'_1(x) + \dots$ converge uniformément sur $[a, b]$ vers une fonction $s(x)$ continue par morceaux, alors la série des fonctions $u_n(x)$ converge uniformément sur $[a, b]$ vers une fonction $g(x)$ lisse par morceaux dont la dérivée $g'(x)$ est égale à $s(x)$ en tout point de continuité de la fonction $s(x)$.

La démonstration se réduit à l'application du théorème 9.77 à la suite des sommes partielles de la série en question.

§ 9.8. Intégration et dérivation par rapport au paramètre

9.81. Soit $f(x, t)$ une fonction numérique d'une variable x parcourant un intervalle $\Delta = \{a \leq x \leq b\}$ et d'une variable t parcourant un espace métrique M . Supposons que cette fonction soit uniformément continue par rapport à l'ensemble des variables x, t , ou bien, ce qui revient au même, uniformément continue sur l'espace métrique $Q = \Delta \times M$ (5.18 et 5.17.a).

T h é o r è m e. Dans les conditions formulées ci-dessus, la fonction

$$\Phi(t) = \int_a^b f(x, t) dx \quad (t \in M)$$

est uniformément continue sur l'espace M .

D é m o n s t r a t i o n. Désignons par

$$\omega_f(\delta) = \sup_{\substack{|x' - x''| \leq \delta \\ \rho(t', t'') \leq \delta}} |f(x', t') - f(x'', t'')|$$

l'oscillation de la fonction $f(x, t)$ sur l'espace Q (5.17.c). Pour un $\varepsilon > 0$ donné, il existe un $\delta_0 > 0$ tel que, si $\delta < \delta_0$, on a $\omega_f(\delta) < \varepsilon$. Soit $\rho(t', t'') \leq \delta < \delta_0$; alors

$$\begin{aligned} |\Phi(t') - \Phi(t'')| &= \left| \int_a^b [f(x, t') - f(x, t'')] dx \right| \leq \\ &\leq \int_a^b |f(x, t') - f(x, t'')| dx \leq \omega_f(\delta)(b-a), \end{aligned}$$

d'où

$$\omega_\Phi(\delta) = \sup_{\rho(t', t'') \leq \delta} |\Phi(t') - \Phi(t'')| \leq \omega_f(\delta)(b-a) < \varepsilon(b-a),$$

ce qui démontre la continuité uniforme de la fonction $\Phi(t)$ sur M .

9.82. Soit $f(x, t)$ une fonction réelle définie et continue sur un rectangle $Q = \{a \leq x \leq b, c \leq t \leq d\}$. Alors

$$\int_a^b \left\{ \int_c^d f(x, t) dt \right\} dx = \int_c^d \left\{ \int_a^b f(x, t) dx \right\} dt. \quad (1)$$

Démonstration. Posons

$$\Phi(t) = \int_a^b f(x, t) dx \quad (c \leq t \leq d),$$

$$F(x) = \int_c^d f(x, t) dt \quad (a \leq x \leq b).$$

En vertu de 9.81, les fonctions $\Phi(t)$ et $F(x)$ sont continues. Rappelons, ensuite, que la différence de l'intégrale d'une fonction continue $G(t)$ sur l'intervalle $[c, d]$ et de sa somme intégrale construite pour une partition Π avec $d(\Pi) \leq \delta$ a pour estimation (9.15.m)

$$\left| \int_c^d G(t) dt - \sum_{j=0}^{m-1} G(\xi_j) \Delta t_j \right| \leq \omega_G(\delta)(d-c). \quad (2)$$

En particulier, pour une partition Π de l'intervalle $[c, d]$ avec $d(\Pi) < \delta$, on a

$$\left| \int_c^d f(x, t) dt - \sum_{j=0}^{m-1} f(x, t_j) \Delta t_j \right| \leq \omega_f(\delta)(d-c).$$

En intégrant cette inégalité par rapport à x on obtient (9.15.b)

$$\begin{aligned} \left| \int_a^b \left\{ \int_c^d f(x, t) dt \right\} dx - \int_a^b \left\{ \sum_{j=0}^{m-1} f(x, t_j) \Delta t_j \right\} dx \right| &\leq \\ &\leq \omega_f(\delta)(d-c)(b-a). \end{aligned} \quad (3)$$

Or, en vertu de la même estimation (2) appliquée à la fonction $\Phi(t)$, on a

$$\begin{aligned} &\left| \int_c^d \Phi(t) dt - \sum_{j=0}^{m-1} \Phi(t_j) \Delta t_j \right| = \\ &= \left| \int_c^d \left\{ \int_a^b f(x, t) dx \right\} dt - \sum_{j=0}^{m-1} \left\{ \int_a^b f(x, t_j) dx \right\} \Delta t_j \right| \leq \\ &\leq \omega_\Phi(\delta)(d-c) \leq \omega_f(\delta)(b-a)(d-c). \end{aligned} \quad (4)$$

Vu que ..

$$\int_a^b \left\{ \sum_{j=0}^{m-1} f(x, t_j) \Delta t_j \right\} dx = \sum_{j=0}^{m-1} \left\{ \int_a^b f(x, t_j) dx \right\} \Delta t_j,$$

il résulte de (3) et (4) que

$$\left| \int_a^b \left\{ \int_c^d f(x, t) dt \right\} dx - \int_c^d \left\{ \int_a^b f(x, t) dx \right\} dt \right| \leq \\ \leq 2\omega_1(\delta) (b-a)(d-c),$$

et comme le second membre peut être rendu aussi petit que l'on veut, l'égalité (1) est valable, ce qu'il fallait démontrer.

9.83. Ex e m p l e. Posons, dans le théorème 9.82, $f(x, t) = x^t$, $x \in [0, 1]$, $t \in [\alpha, \beta]$ ($\alpha \geq 0$, $\beta > 0$); nous obtenons

$$\int_0^1 \left\{ \int_{\alpha}^{\beta} x^t dt \right\} dx = \int_{\alpha}^{\beta} \left\{ \int_0^1 x^t dx \right\} dt.$$

En calculant les intégrales entre accolades trouvons la valeur de l'intégrale définie

$$\int_0^1 \frac{x^{\beta} - x^{\alpha}}{\ln x} dx = \ln \frac{1+\beta}{1+\alpha}.$$

Il serait difficile de le faire par des procédés usuels décrits dans §§ 9.4-9.5, l'intégrale indéfinie de $\frac{x^{\beta} - x^{\alpha}}{\ln x}$ ne s'exprimant pas par les fonctions élémentaires.

9.84. T h é o r è m e. Si une fonction $f(x, t)$ est définie pour $a \leq x \leq b$, $c \leq t \leq d$ et a sa dérivée continue par rapport à t dans le voisinage d'un point $t_0 \in (c, d)$:

$$f_i(x, t) = \lim_{\Delta t \rightarrow 0} \frac{f(x, t + \Delta t) - f(x, t)}{\Delta t}, \quad |t - t_0| < \delta,$$

alors on a l'égalité

$$\left[\frac{d}{dt} \int_a^b f(x, t) dx \right]_{t=t_0} = \int_a^b f_i(x, t_0) dx \quad (1)$$

pour tout $t_0 \in [c, d]$.

D é m o n s t r a t i o n. La fonction $f_i(x, t)$ étant continue dans le rectangle $a \leq x \leq b$, $t_0 - \delta \leq t \leq t_0 + \delta$, on peut y appliquer le théorème 9.82; en se servant de la formule de Newton-Leibniz (9.32)

on obtient ensuite

$$\begin{aligned} \int_{t_0}^t \left\{ \int_a^b f'_1(x, t) dx \right\} dt &= \int_a^b \left\{ \int_{t_0}^t f'_1(x, t) dt \right\} dx = \\ &= \int_a^b f(x, t) dx - \int_a^b f(x, t_0) dx \end{aligned}$$

pour tout $t \in [t_0 - \delta, t_0 + \delta]$.

Le premier membre a la dérivée en t égale à la fonction sous le signe somme (9.31). Il va de soi que le second membre a la même dérivée. Or, le deuxième terme du second membre ne dépend pas de t , donc sa dérivée en t est nulle. En égalant les dérivées des deux membres et en posant $t = t_0$, on aboutit à (1). Le théorème est démontré.

9.85. E x e m p l e. L'intégrale

$$I(t) = \int_0^{\pi/2} \ln(t^2 - \sin^2 x) dx \quad (t > 1)$$

ne peut être calculée par les méthodes de § 9.4. Tout de même, d'après le théorème 9.84 dont l'hypothèse est, évidemment, remplie, on a

$$I'(t) = \int_0^{\pi/2} \frac{2t dx}{t^2 - \sin^2 x} = \frac{\pi}{\sqrt{t^2 - 1}}$$

(ici on peut déjà trouver une primitive à l'aide de 9.43.b). En intégrant cette égalité par rapport à t on retrouve la valeur de l'intégrale

$$I(t) = \int \frac{\pi}{\sqrt{t^2 - 1}} dt = \pi \ln(t + \sqrt{t^2 - 1}) + C.$$

Afin de déterminer la valeur de C , on passe à la limite pour $t \rightarrow \infty$ ou $\tau = \frac{1}{t} \rightarrow 0$ dans l'égalité

$$\begin{aligned} C &= \int_0^{\pi/2} \ln(t^2 - \sin^2 x) dx - \pi \ln(t + \sqrt{t^2 - 1}) = \\ &= \int_0^{\pi/2} \ln\left(1 - \frac{\sin^2 x}{t^2}\right) dx - \pi \ln \frac{t + \sqrt{t^2 - 1}}{t} \end{aligned}$$

valable pour tous les $t > 1$. Notons que la fonction

$$\ln(1 - \tau^2 \sin^2 x)$$

est définie et continue dans le rectangle $0 \leq x \leq \frac{\pi}{2}$, $0 \leq \tau \leq \tau_0 < 1$, donc la fonction

$$\Phi(\tau) = \int_0^{\pi/2} \ln(1 - \tau^2 \sin^2 x) dx$$

est continue pour $0 \leq \tau \leq \tau_0$ en vertu de 9.81. Par conséquent,

$$0 = \Phi(0) = \lim_{\tau \rightarrow 0} \Phi(\tau) = \lim_{\tau \rightarrow 0} \int_0^{\pi/2} \ln(1 - \tau^2 \sin^2 x) dx.$$

D'autre part,

$$\lim_{t \rightarrow \infty} \pi \ln \frac{t + \sqrt{t^2 - 1}}{t} = \pi \lim_{\tau \rightarrow 0} \ln(1 + \sqrt{1 - \tau^2}) = \pi \ln 2.$$

Définitivement, on a $C = -\pi \ln 2$ et

$$I(t) = \pi \ln \frac{t + \sqrt{t^2 - 1}}{2}.$$

9.86. Voici un cas un peu plus délicat où non seulement la fonction à intégrer mais aussi les limites d'intégration dépendent du paramètre t .

a. **Théorème.** Soit $f(x, t)$ une fonction continue pour $a \leq x \leq b$, $c \leq t \leq d$ dont la dérivée $f_t(x, t)$ existe et est continue pour $a \leq x \leq b$, $|t - t_0| \leq \delta$. Si une fonction $\varphi(t)$ est définie, continue et dérivable pour $|t - t_0| < \delta$, à valeurs dans l'intervalle $a \leq x \leq b$, la fonction

$$\Phi(t) = \int_0^{\varphi(t)} f(x, t) dx$$

est dérivable pour $t = t_0$, et

$$\Phi'(t_0) = \int_a^{\varphi(t_0)} f_t(x, t_0) dx + f(\varphi(t_0), t_0) \cdot \varphi'(t_0). \quad (1)$$

Démonstration. Posons $\Phi(t) = \Phi_1(t) + \Phi_2(t)$, où

$$\Phi_1(t) = \int_a^{\varphi(t_0)} f(x, t) dx, \quad \Phi_2(t) = \int_{\varphi(t_0)}^{\varphi(t)} f(x, t) dx.$$

D'après le théorème 9.84, la fonction $\Phi_1(t)$ a, pour $t = t_0$, la dérivée

$$\Phi_1'(t_0) = \int_a^{\varphi(t_0)} f_t(x, t_0) dx. \quad (2)$$

Trouvons la dérivée de $\Phi_2(t)$. Nous avons, en vertu du théorème de la moyenne 9.15.h,

$$\begin{aligned}\frac{\Phi_2(t_0 + \Delta t) - \Phi_2(t_0)}{\Delta t} &= \frac{1}{\Delta t} \int_{\varphi(t_0)}^{\varphi(t_0 + \Delta t)} f(x, t_0 + \Delta t) dx = \\ &= \frac{\varphi(t_0 + \Delta t) - \varphi(t_0)}{\Delta t} f(\theta, t_0 + \Delta t); \end{aligned}$$

le nombre θ est compris entre $\varphi(t_0)$ et $\varphi(t_0 + \Delta t)$. En vertu des suppositions relatives aux fonctions $f(x, t)$ et $\varphi(t)$, l'expression obtenue tend, pour $\Delta t \rightarrow 0$, vers la limite

$$\Phi'_2(t_0) = \varphi'(t_0) \cdot f(\varphi(t_0), t_0). \quad (3)$$

Définitivement, (1) résulte de (2) et (3), et le théorème est démontré.

b. Soit, en particulier $c = a$, $d = b$, $\varphi(t) \equiv t$. On a la formule

$$\frac{d}{dt} \int_a^t f(x, t) dx \Big|_{t=t_0} = \int_a^{t_0} f_t(x, t_0) dx + f(t_0, t_0) \quad (4)$$

valable dans l'hypothèse que $f(x, t)$ soit continue pour $a \leq x \leq b$, $a \leq t \leq b$ et que $f_t(x, t)$ existe et soit continue au voisinage du point $t = t_0$.

9.87. Ex e m p l e. Trouvons la dérivée par rapport à x de la fonction

$$I_n(x) = \frac{1}{(n-1)!} \int_a^x (x-u)^{n-1} f(u) du \quad (n=1, 2, \dots), \quad (1)$$

où $f(x)$ est une fonction continue pour $a \leq x \leq b$.

D'après la formule 9.86 (4), où t est remplacé par x et x par u , on a

$$\begin{aligned}I'_n(x) &= \frac{1}{(n-1)!} \int_a^x (n-1)(x-u)^{n-2} f(u) du + \frac{(x-x)^{n-1}}{(n-1)!} f(x) = \\ &= I_{n-1}(x). \end{aligned}$$

De plus, on a, évidemment, $I_n(a) = 0$, de sorte que

$$I_n(x) = \int_a^x I'_n(\xi) d\xi = \int_a^x I_{n-1}(\xi) d\xi.$$

Notons encore que $I_1(x) = \int_a^x f(u) du$.

Il en découle que $I_n(x)$ est le résultat de n intégrations successives de la fonction $f(x)$ de a à x , autrement dit, de l'opération inverse

par rapport à n dérivations successives. La formule (1) met le résultat de cette opération sous la forme d'une intégrale simple avec paramètre.

§ 9.9. Intégrales curvilignes

9.91. Intégrale curviligne de Stieltjes. Considérons dans l'espace n -dimensionnel R_n une courbe L dont la représentation paramétrique est $x_j = x_j(t)$ ($j = 1, 2, \dots, n$), où les fonctions $x_j(t)$ sont définies et continues sur un intervalle $[a, b]$ et y possèdent leurs dérivées continues (ou continues par morceaux). La courbe parcourue dans le sens des t croissants de a à b est appelée *chemin*. D'après 9.63.g, si $\sum_1^n [x_j(t+0)]^2$ et $\sum_1^n [x_j(t-0)]^2$ ne s'annulent pas, on peut choisir pour paramètre sur le chemin L la longueur d'arc s comptée à partir d'un point fixe A .

Considérons une partie du chemin L entre le point A ($s = 0$) et un point B ($s = S$). Désignons par Π la partition de l'arc AB par des points successifs $A = M_0, M_1, \dots, M_p = B$. Les valeurs correspondantes du paramètre s , notamment

$$s(A) = 0 \leq s(M_1) = s_1 \leq \dots \leq s(B) = s_p = S,$$

forment une partition de l'intervalle $[0, S]$ de l'axe réel. Posons comme d'habitude

$$d(\Pi) = \max (s_{j+1} - s_j) = \max \Delta s_j \quad (j = 1, \dots, p-1)$$

Soient $f(M)$ et $g(M)$ deux fonctions réelles données sur le chemin $\widehat{AB}$. Formons la somme

$$\sum_{j=0}^{p-1} f(M_j) [g(M_{j+1}) - g(M_j)], \quad (1)$$

où chaque point M_j appartient à l'arc $\widehat{M_j M_{j+1}}$, de sorte que $s(M_j) = s_j \in [s_j, s_{j+1}]$. La limite de la somme (1) pour $d(\Pi) \rightarrow 0$, si elle existe, est appelée *intégrale curviligne de Stieltjes de la fonction $f(M)$*

selon la fonction $g(M)$ prise le long de l'arc $\widehat{AB}$ et désignée par

$$\int_{\widehat{AB}} f(M) dg(M). \quad (2)$$

Par exemple, on a

$$\int_{\widehat{AB}} dg(M) = g(B) - g(A).$$

Si la fonction $g(M)$ est constante sur l'arc $\widehat{AB}$, alors, évidemment,

$$\int_{\widehat{AB}} f(M) dg(M) = 0$$

Si l'une quelconque des intégrales ci-dessous existe, alors l'autre existe aussi, et l'on a

$$\int_{\widehat{AB}} f(M) dg(M) = - \int_{\widehat{BA}} f(M) dg(M).$$

Lorsque le chemin $\widehat{AB}$ est donné par ses équations paramétriques $x_k = x_k(t)$ ($k = 1, 2, \dots, n$), $a \leq t \leq b$, on peut dire que $f(M)$ et $g(M)$ sont les fonctions du paramètre t : $f = f(t)$, $g = g(t)$; alors la somme intégrale (1) se met sous la forme

$$\sum_{j=0}^{p-1} f(t_j) [g(t_{j+1}) - g(t_j)], \quad t_j \leq t_{j+1}. \quad (3)$$

Lorsqu'on morcelle indéfiniment une partition du chemin $\widehat{AB}$, il en est de même de la partition correspondante de l'intervalle $[a, b]$ de variation de t , puisque t est une fonction dérivable de la longueur d'arc s ; on a donc, d'après la formule de Lagrange,

$$\Delta t_j = t'(\xi_j) \Delta s_j, \quad s_j \leq \xi_j \leq s_{j+1}$$

Inversement, pour la même raison, à un morcellement illimité d'une partition de l'intervalle $[a, b]$ correspond un morcellement illimité de la partition du chemin $\widehat{AB}$. Donc, la convergence de la somme (1) vers sa limite est équivalente à la convergence de la somme (3), et le problème d'existence de l'intégrale curviligne de Stieltjes (2) se réduit au problème d'existence de l'intégrale usuelle de Stieltjes (9.56):

$$\int_a^b f(t) dg(t) \quad (4)$$

qui, pour une fonction lisse $g(t)$, est une intégrale de Riemann 9.56 (4).

Par conséquent, les formules suivantes, analogues à 9.56 (5)-(8), sont valables:

$$\begin{aligned} & \int_{\widehat{AB}} [\alpha_1 f_1(M) + \alpha_2 f_2(M)] dg(M) = \\ &= \alpha_1 \int_{\widehat{AB}} f_1(M) dg(M) + \alpha_2 \int_{\widehat{AB}} f_2(M) dg(M); \end{aligned} \quad (5)$$

$$\int_{\widehat{AB}} f(M) d[\beta_1 g_1(M) + \beta_2 g_2(M)] =$$

$$= \beta_1 \int_{\widehat{AB}} f(M) dg_1(M) + \beta_2 \int_{\widehat{AB}} f(M) dg_2(M); \quad (6)$$

$$\int_{\widehat{AB}} f(M) dg(M) = \int_{\widehat{AC}} f(M) dg(M) + \int_{\widehat{CB}} f(M) dg(M) \quad (\widehat{AB} \supset C); \quad (7)$$

$$\int_{\widehat{AB}} f(M) dg(M) = f(M) g(M) \Big|_A^B - \int_{\widehat{AB}} g(M) df(M), \quad (8)$$

dès que les intégrales dans les seconds membres existent.

On a souvent affaire à une somme d'intégrales curvilignes prises le long d'un même chemin

$$\int_{\widehat{AB}} f_1(M) dg_1(M) + \dots + \int_{\widehat{AB}} f_p(M) dg_p(M).$$

Sa notation plus brève est

$$\int_{\widehat{AB}} f_1(M) dg_1(M) + \dots + f_p(M) dg_p(M) =$$

$$= \int_{\widehat{AB}} f_1 dg_1 + \dots + f_p dg_p$$

(sans parenthèses). En supposant que les fonctions $f_j(t)$ soient continues et les fonctions $g_j(t)$ lisses par morceaux, la somme est calculée d'après la règle 9.56 (4)

$$\int_a^b [f_1(t) g'_1(t) + \dots + f_p(t) g'_p(t)] dt. \quad (9)$$

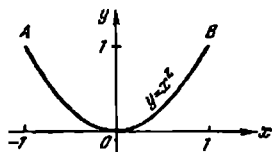


Fig. 9.42.

9.92. Exemples. Calculons les intégrales curvilignes

$$I_1 = \int_{\widehat{AB}} (x^2 - y^2) dx, \quad I_2 = \int_{\widehat{AB}} (x^2 - y^2) dy,$$

$\widehat{AB}$ étant l'arc de la parabole $y = x^2$ entre les points $A = (-1, 1)$ et $B = (1, 1)$ (fig. 9.42). Dans le premier cas, on peut choisir x pour paramètre, et l'on a

$$I_1 = \int_{-1}^1 (x^2 - x^4) dx = \left(\frac{x^3}{3} - \frac{x^5}{5} \right) \Big|_{-1}^1 = \frac{2}{3} - \frac{2}{5} = \frac{4}{15}.$$

Dans le second, il est impossible de choisir y pour paramètre car, à une seule valeur de y correspondent deux points de l'arc. Pour cette raison, c'est toujours x qui sera le paramètre; on aura, d'après 9.56 (4):

$$I_2 = \int_{-1}^1 (x^2 - x^4) 2x dx = \left[\frac{2x^4}{4} - \frac{2x^5}{5} \right]_{-1}^1 = 0.$$

9.93. Intégrale le long d'un contour fermé. Si l'extrémité B d'un chemin $\widehat{AB}$ se confond avec l'origine A , le chemin $\widehat{AB}$ s'appelle *contour fermé*. L'intégrale le long d'un contour fermé L est souvent désignée par

$$\oint_L f dg. \quad (1)$$

La valeur de l'intégrale (1) ne dépend pas du choix de A car, pour n'importe quelle autre origine A' , on a

$$\int_{\widehat{AA'A}} f dg = \int_{\widehat{AA'}} f dg + \int_{\widehat{A'A}} f dg = \int_{\widehat{A'A}} f dg + \int_{\widehat{AA'}} f dg = \int_{\widehat{A'A A'}} f dg$$

d'après 9.91 (7).

Néanmoins, l'intégrale (1) dépend du sens de parcours du contour; en effet, en inversant le sens de parcours, toutes les différences $g(M_{j+1}) - g(M_j)$ changent de signe, donc l'intégrale (1) change de signe elle aussi. Le sens de parcours est parfois indiqué par une flèche

$$\oint_L f dg = - \oint_L f dg. \quad (2)$$

D'ailleurs, cette indication n'est utile que dans le cas du plan ($n = 2$).

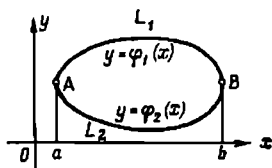


Fig. 9.43.

9.94. Calcul de l'aire d'une figure plane à l'aide de l'intégrale curviligne. Soit une figure plane (fig. 9.43) limitée par les courbes $y = \varphi_1(x)$ et $y = \varphi_2(x)$. Son aire est

$$S = \int_a^b \varphi_1(x) dx - \int_a^b \varphi_2(x) dx. \quad (1)$$

Notons que, par définition, les deux intégrales peuvent être considérées comme curvilignes, de sorte que

$$S = \int_{L_1} y dx - \int_{L_2} y dx.$$

On peut noter alors

$$S = \int_{L_1} y \, dx + \int_{-L_2} y \, dx \quad (-L_2 \text{ est } L_2 \text{ parcouru de } B \text{ à } A)$$

et ajouter ensuite les intégrales (nulles) prises le long des segments verticaux (s'il en existe un ou deux); la somme de toutes ces intégrales représente l'intégrale prise le long du contour fermé:

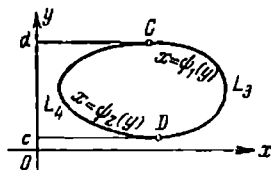


Fig. 9.44.

$$S = \oint_L y \, dx = - \oint_L y \, dx. \quad (2)$$

Ainsi, l'aire d'une figure s'exprime par une intégrale curviligne prise le long de son contour L .

On peut établir une formule plus symétrique. Notamment, en faisant changer de rôles les coordonnées x et y , on trouve (fig. 9.44):

$$\begin{aligned} S &= \int_c^d \psi_1(y) \, dy - \int_c^d \psi_2(y) \, dy = \int_{L_3} x \, dy - \int_{L_1} x \, dy = \\ &= \int_{L_3} x \, dy + \int_{-L_4} x \, dy = \oint_L x \, dy. \end{aligned} \quad (3)$$

L'addition de (3) et (2) et la multiplication par $1/2$ fournissent

$$S = \frac{1}{2} \oint_L x \, dy - y \, dx \quad (4)$$

Le paramètre étant convenablement choisi, la formule (4) peut s'avérer plus commode que (1). Par exemple, pour une ellipse $x = a \cos t$, $y = b \sin t$, on obtient directement

$$S = \frac{1}{2} \oint_L x \, dy - y \, dx = \frac{1}{2} \int_0^{2\pi} ab [\cos^2 t + \sin^2 t] \, dt = \pi ab.$$

est bon de comparer la formule (4) à la formule

$$\int_A^B x \, dy + y \, dx = \int_A^B d(xy) = xy \Big|_A^B = x(B)y(B) - x(A)y(A);$$

pour un contour fermé Γ , on a

$$\oint_{\Gamma} x \, dy + y \, dx = 0.$$

D'une manière analogue, pour tout contour fermé Γ et pour deux fonctions continues $f(x)$, $\varphi(y)$, on a

$$\oint_{\Gamma} f(x) dx = 0; \quad \oint_{\Gamma} \varphi(y) dy = 0.$$

9.95. Approximation d'une intégrale prise le long d'une courbe par une intégrale prise le long d'une ligne polygonale. Ici on se borne aux intégrales curvillignes de la forme

$$\int_L f_1(M) dx_1 + \dots + f_n(M) dx_n \quad (1)$$

en supposant que L soit un chemin dans l'espace n -dimensionnel dont la représentation paramétrique est $x_j = x_j(t)$, $a \leq t \leq b$, $j = 1, \dots, n$, les fonctions $x_j(t)$ étant continues et telles que $\sum_{j=1}^n [x_j(t)]^2 > 0$. (D'ailleurs, on peut les supposer continues par morceaux; alors on partage le chemin L en un nombre fini d'arcs avec $x_j(t)$ continues.)

Les intégrales de la forme (1) admettent l'estimation suivante par la longueur $s(L)$ du chemin L :

$$\left| \int_L f_1(M) dx_1 + \dots + f_n(M) dx_n \right| \leq \max_{M \in L} \sqrt{\sum_{k=1}^n f_k^2(M)} \cdot s(L). \quad (2)$$

En effet, si l'on considère une somme intégrale composée pour une partition Π , on obtient en appliquant l'inégalité 2.68(4):

$$\begin{aligned} \left| \sum_{j=0}^{p-1} \sum_{k=1}^n f_k(M_k) \Delta_j x_k \right| &\leq \sum_{j=0}^{p-1} \sqrt{\sum_{k=1}^n f_k^2(M_k)} \times \\ &\times \sqrt{\sum_{k=1}^n (\Delta_j x_k)^2} \leq \max_M \sqrt{\sum_{k=1}^n f_k^2(M)} \cdot \sum_{j=0}^{p-1} \Delta h_j, \end{aligned}$$

où Δh_j est la longueur de la corde $M_j M_{j+1}$. La somme des longueurs de ces cordes étant la longueur de la ligne polygonale $M_0 M_1 \dots M_p$, elle ne dépasse pas la longueur $s(L)$ de l'arc L même. En passant à la limite on obtient (2).

Supposons ensuite que les fonctions $f_1(M), \dots, f_n(M)$ soient définies et continues non seulement sur la courbe L , mais aussi dans un voisinage de L qui contient toutes les lignes polygonales,

à côtés assez petits, inscrites dans L . Soient une partition $\Pi = \{a = t_0 \leq t_1 \leq \dots \leq t_p = b\}$ de l'intervalle $[a, b]$ de variation du paramètre t et L_Π une ligne polygonale à sommets successifs $A = M_0, M_1, \dots, M_p = B$ sur la courbe L correspondant aux valeurs $t_0, t_1, \dots, t_p$ du paramètre t . On a

$$\begin{aligned} \lim_{d(\Pi) \rightarrow 0} \int_{L_\Pi} f_1(M) dx_1 + \dots + f_n(M) dx_n &= \\ &= \int_L f_1(M) dx_1 + \dots + f_n(M) dx_n. \end{aligned} \quad (3)$$

Il en résulte que, les intégrales le long de toutes les lignes polygonales d'un domaine de l'espace R_n étant données, on peut calculer l'intégrale le long de toute courbe contenue dans ce domaine.

Il suffit de démontrer la relation (3) pour $n = 1$:

$$\lim_{d(\Pi) \rightarrow 0} \int_{L_\Pi} f_1(M) dx_1 = \int_L f_1(M) dx_1. \quad (4)$$

En faisant la démonstration nous omettons l'indice 1. Soit un $\epsilon > 0$; choisissons un $\delta > 0$ tel que

$$\left| \int_L f(M) dx - \sum_{j=0}^{p-1} f(M_j) \Delta x_j \right| < \epsilon \quad (5)$$

pour $d(\Pi) < \delta$ et que $\rho(M, M') < \delta$ implique $|f(M) - f(M')| < \epsilon$.

Considérons d'abord un seul côté de la ligne polygonale L_Π , par exemple $M_j M_{j+1}$. On a, en se servant de l'estimation (2):

$$\int_{M_j M_{j+1}} f(M_j) dx = f(M_j) \int_{M_j M_{j+1}} dx = f(M_j) \Delta x_j$$

et

$$\begin{aligned} & \left| \int_{M_j M_{j+1}} f(M) dx - f(M_j) \Delta x_j \right| = \\ &= \left| \int_{M_j M_{j+1}} f(M) dx - \int_{M_j M_{j+1}} f(M_j) dx \right| = \\ &= \left| \int_{M_j M_{j+1}} [f(M) - f(M_j)] dx \right| \leq \epsilon s(M_j M_{j+1}). \end{aligned}$$

En sommant suivant tous les côtés, on obtient

$$\left| \int_{L_\Pi} f(M) dx - \sum_{j=0}^{p-1} f(M_j) \Delta x_j \right| \leq \epsilon \sum_{j=0}^{p-1} s(M_j M_{j+1}) = \epsilon s(L_\Pi). \quad (6)$$

Les longueurs $s(L_{\Pi})$ des lignes polygonales L_{Π} étant bornées (elles ont pour limite la longueur de l'arc L), les inégalités (5) et (6) impliquent la relation (4). Par conséquent, (3) est établi.

Exercices

1. Si $f(x)$ est une fonction périodique sur $(-\infty, \infty)$, alors la fonction

$$F(x) = \int_a^x f(\xi) d\xi$$

est décomposable en somme d'une fonction périodique et d'une fonction du premier degré.

2. Si $f(x)$ et $\varphi(x) \geq 0$ sont bornées et intégrables sur $[a, b]$ et si $\varphi(x)$ ne décroît pas, alors il existe un point $\xi \in [a, b]$ tel que

$$\int_a^b f(x) \varphi(x) dx = \varphi(b-0) \int_a^b f(x) dx.$$

Si $\varphi(x) \geq 0$ ne croît pas, alors il existe un point $\eta \in [a, b]$ tel que

$$\int_a^b f(x) \varphi(x) dx = \varphi(a+0) \int_a^b f(x) dx.$$

3. Si $f(x)$ et $\varphi(x)$ sont bornées et intégrables sur $[a, b]$ et si $\varphi(x)$ est monotone, alors il existe un point $\xi \in [a, b]$ tel que

$$\int_a^b f(x) \varphi(x) dx = \varphi(a+0) \int_a^{\xi} f(x) dx + \varphi(b-0) \int_{\xi}^b f(x) dx$$

(« second théorème de la moyenne »).

4. Démontrer la formule d'intégration multiple par parties:

$$\int_a^b uv^{(n+1)}(x) dx = [uv^{(n)} - u'v^{(n-1)} + \dots + (-1)^n u^{(n)}v] \Big|_a^b + (-1)^{n+1} \int_a^b u^{(n+1)}v dx$$

dans les suppositions convenables sur la dérivabilité des fonctions u et v .

5. Pour une fonction bornée $f(x)$ et pour un intervalle $[\alpha, \beta]$, posons

$$\omega[f; \alpha, \beta] = \sup_{x', x'' \in [\alpha, \beta]} |f(x') - f(x'')|.$$

Démontrer que la fonction $f(x)$ est intégrable sur $[a, b]$ si, et seulement si,

$$\lim_{d(\Pi) \rightarrow 0} \sum_{j=0}^{n-1} \omega[f; x_j, x_{j+1}] (x_{j+1} - x_j) = 0,$$

où

$\Pi = \{a = x_0 < x_1 < \dots < x_n = b\}$, $d(\Pi) = \max_j (x_{j+1} - x_j)$ (critère de Riemann).

6. Démontrer qu'une fonction monotone et bornée est intégrable.
 7. Si une fonction $f(x)$ est intégrable, alors la fonction $|f(x)|$ l'est aussi.
 8. On appelle *oscillation d'une fonction $f(x)$ en un point c* le nombre

$$\omega[f; c] = \inf_{\alpha, \beta} \omega[f; \alpha, \beta], \quad \alpha < c < \beta.$$

Démontrer que pour qu'une fonction bornée $f(x)$ soit intégrable sur $[a, b]$, il faut et il suffit que, pour n'importe quels $\varepsilon > 0$ et $\delta > 0$, l'ensemble des points $c \in [a, b]$ avec $\omega[f; c] \geq \varepsilon$ appartienne à un nombre fini d'intervalles dont la somme des longueurs est $\leq \delta$ (*critère de Du Bois-Reymond*).

9. Démontrer que pour qu'une fonction bornée $f(x)$ soit intégrable sur $[a, b]$, il faut et il suffit que, pour tout $\delta > 0$, l'ensemble des points de discontinuité de $f(x)$ soit inclus dans une famille finie ou dénombrable d'intervalles dont la somme des longueurs est $\leq \delta$ (*critère de Lebesgue*).

10. Soit $f(x)$ une fonction intégrable. Démontrer que la fonction $\frac{1}{f(x)}$ est intégrable elle aussi si elle est bornée.

11. En intégrant l'inégalité double $\sin^{2n+1} x \leq \sin^{2n} x \leq \sin^{2n-1} x$ sur l'intervalle $0 \leq x \leq \frac{\pi}{2}$ et en se servant des formules 9.51 (1), (2), établir la *formule de Wallis* (1656)

$$\frac{\pi}{2} = \lim_{n \rightarrow \infty} \frac{2 \cdot 2 \cdot 4 \cdot 4 \dots 2n \cdot 2n}{1 \cdot 3 \cdot 3 \cdot 5 \dots (2n-1)(2n+1)} = \prod_{n=1}^{\infty} \frac{1}{1 - \frac{1}{4n^2}}.$$

12. Soit $f(x)$ une fonction continue sur un intervalle $[A, B]$ qui contient deux points a et b (disposés d'une façon quelconque). Soit, ensuite $\Pi = (a = x_0, x_1, \dots, x_n = b)$ un ensemble de points de l'intervalle $[A, B]$ (disposés dans n'importe quel ordre). Formons une « somme intégrale généralisée »

$$s_{\Pi}(f) = \sum_{j=0}^{n-1} f(\xi_j)(x_{j+1} - x_j), \quad (1)$$

où ξ_j est un point quelconque entre x_j et x_{j+1} . Démontrer que, pour $\sum_{j=0}^{n-1} |x_{j+1} - x_j| < C$ et $d(\Pi) = \max |x_{j+1} - x_j| \rightarrow 0$, la somme (1) tend vers une limite et que

$$\lim_{d(\Pi) \rightarrow 0} s_{\Pi}(f) = \int_a^b f(x) dx. \quad (2)$$

13. Si, dans l'exercice 12, on supprime l'hypothèse $\sum_{j=0}^{n-1} |x_{j+1} - x_j| < C$ ou bien on ne suppose que la continuité par morceaux de la fonction $f(x)$, alors la somme (1) peut s'avérer divergente.

14. (*Définition axiomatique de l'intégrale*.) Supposons que, à toute fonction $f(x)$ continue par morceaux sur un intervalle $[a, b]$ et à tout intervalle $[\alpha, \beta] \subset [a, b]$ on fasse correspondre un nombre $I_{\alpha}^{\beta}(f)$ satisfaisant aux conditions suivantes:

- $I_{\alpha}^{\beta}(kf) = k I_{\alpha}^{\beta}(f)$, pour toute constante k ;
- $I_{\alpha}^{\beta}(f) + I_{\alpha}^{\beta}(g) = I_{\alpha}^{\beta}(f+g)$, quelles que soient les fonctions $f(x)$ et $g(x)$ continues par morceaux;

$$3) I_{\alpha}^{\beta}(1) = \beta - \alpha ;$$

$$4) I_{\alpha}^{\beta}(f) = I_{\alpha}^{\gamma}(f) + I_{\gamma}^{\beta}(f), \text{ quels que soient } \alpha < \gamma < \beta ;$$

$$5) |I_{\alpha}^{\beta}(f)| \leq C \sup |f(x)| \text{ avec une constante fixe } C.$$

Alors, quelle que soit une fonction $f(x)$ continue par morceaux, on a

$$I_{\alpha}^{\beta}(f) = \int_{\alpha}^{\beta} f(x) dx. \quad (3)$$

15. La convergence de la série 9.74 (6) étant trop lente, elle est peu appropriée au calcul du nombre π . Montrer que le développement suivant, qui représente un moyen efficace de calcul de π , est valable :

$$\frac{\pi}{6} = \frac{1}{\sqrt{3}} \left(1 - \frac{1}{3 \cdot 3} + \frac{1}{3^2 \cdot 5} - \frac{1}{3^3 \cdot 7} + \dots \right) = 0,523599 \dots$$

Historique

Les problèmes géométriques qui, dans le langage moderne, sont notés $\int_{\alpha}^{\beta} x dx$

ou $\int_{\alpha}^{\beta} x^2 dx$ sont considérés déjà dans l'Antiquité (Archimède). Le calcul intégral

dans notre sens avec ses applications géométriques, mécaniques et physiques est l'œuvre du XVII^e siècle et est essentiellement dû à Newton et Leibniz. Newton appelle fluxion une dérivée et fluente une primitive. Les symboles d et $\int$ ainsi que les règles de calcul des intégrales indéfinies appartiennent à Leibniz. Les intégrales curvilignes apparaissent pour la première fois chez Clairaut (1743).

La définition rigoureuse de l'intégrale en tant que limite des sommes intégrales est donnée par Cauchy (1821), après quoi la position du problème d'existence de l'intégrale des fonctions de telle ou telle classe devient, enfin, possible. Cauchy propose une démonstration de l'existence de l'intégrale d'une fonction continue qui s'avère pourtant incorrecte, faute de notion de continuité uniforme. La première démonstration correcte de l'existence de l'intégrale d'une fonction continue est donnée par Darboux en 1875. Certaines conditions nécessaires et suffisantes de l'intégrabilité d'une fonction (discontinue) sont successivement données par Riemann, Du Bois-Reymond et Lebesgue dans la seconde moitié du XIX^e siècle. En 1894 Stieltjes introduit une nouvelle notion d'intégrale en rapport avec certains problèmes spéciaux ; au XX^e siècle, cette notion est largement appliquée à des questions d'ordre général. En 1902, Lebesgue formule une notion d'intégrale plus générale que les précédentes ; dans les mathématiques modernes, elle joue un rôle décisif : premièrement, l'ensemble des fonctions intégrables au sens de Lebesgue peut être considéré comme espace normé complet (cf. Troisième partie, chapitre 12), et, deuxièmement, il devient possible de décrire la classe de toutes les primitives d'après leurs propriétés intrinsèques. Cf. [14].

Fonctions analytiques

Ainsi, nous comparons une fonction analytique à un organisme dont la propriété caractéristique est de réagir en bloc à toute action sur sa partie quelconque.

G. Polya et G. Szegő

Problèmes et théorèmes de l'analyse (1924)

§ 10.1. Définitions et exemples

10.11. Dans le présent chapitre, nous étudierons les fonctions $f(z)$ définies sur un ensemble E du plan C_z d'une variable complexe z et à valeurs dans le plan C_w d'une variable complexe w .

Définissons la dérivée par rapport à un ensemble $E \subset C_z$ d'une fonction $f(z)$ en un point non isolé $z_0 \in E$.

a. Définition. Un nombre A s'appelle *dérivée par rapport à un ensemble* $E \subset C_z$ d'une fonction $w = f(z)$ en un point non isolé $z_0 \in E$ si, pour un $\varepsilon > 0$ donné, il existe un $\delta > 0$ tel que les relations $|z - z_0| < \delta$, $z \in E$, $z \neq z_0$ impliquent l'inégalité

$\left| A - \frac{f(z) - f(z_0)}{z - z_0} \right| < \varepsilon$. Le nombre A peut alors être désigné par $f'_E(z_0)$. Il est, évidemment, la limite du rapport

$$\frac{f(z) - f(z_0)}{z - z_0} \quad (1)$$

suivant la direction $z \xrightarrow{E} z_0$ qui est définie par la famille d'intersections de tout cercle $|z - z_0| \leq \rho$ dépourvu du point au centre z_0 avec l'ensemble E . Cette définition rappelle celle de la dérivée d'une fonction réelle dans le domaine réel (7.11), qui en est un cas particulier: pour avoir 7.11, il faut choisir pour E un intervalle de l'axe réel contenant le point $z_0 = x_0$. Tout de même, à côté d'une analogie apparente des dérivées dans les domaines réel et complexe, il y a aussi certaines distinctions essentielles qui seront précisées dans la suite.

b. Si F est une partie de l'ensemble E et z_0 un point non isolé de l'ensemble F (donc aussi de E), alors l'existence de la dérivée $f'_E(z)$ implique l'existence de $f'_F(z)$ ainsi que l'égalité

$$f'_F(z) = f'_E(z).$$

Donc, la *dérivabilité par rapport au plus grand de deux ensembles* implique la *dérivabilité par rapport au plus petit*. La réciproque est manifestement fausse, ce qu'on verra sur de simples exemples.

c. Si une fonction $w = f(z)$ est dérivable en un point $z_0 \in E$ par rapport à un ensemble E , alors le rapport (1) est majoré, pour $z \rightarrow z_0$, par une constante C ; on a donc, pour $z \in E$ et $|z - z_0|$ assez petits :

$$|f(z) - f(z_0)| < C |z - z_0|.$$

Par conséquent, la fonction $f(z)$ est continue sur E pour $z = z_0$.

d. **Définition.** Une fonction $w = f(z)$ ayant sa dérivée par rapport à un ensemble E en tout point de E s'appelle *dérivable dans E* .

e. **Définition.** Conformément à la définition 3.21, un ensemble G dans le plan C_z qui, avec tout son point z_0 , contient un cercle $|z - z_0| < r$ (r peut dépendre de z_0) s'appelle *ensemble ouvert* ou *domaine*.

Un domaine G est dit *connexe* s'il est possible de joindre ses deux points quelconques z_0 et z_1 par une courbe lisse par morceaux contenue entièrement dans G .

f. **Définition.** Si G est un ensemble ouvert dans C_z , alors une fonction $f(z)$ dérivable dans G est dite *analytique* dans G (on dit aussi *holomorphe* ou *régulière*). Sa dérivée par rapport à G est désignée par $f'(z)$ (sans indiquer l'ensemble G). En tout point $z_0 \in G$, cette fonction est dérivable par rapport à n'importe quel ensemble E dont z_0 est un point non isolé, et l'on a

$$f'_E(z_0) = f'(z_0).$$

g. **Définition.** Une fonction $f(z)$ est dite *analytique en un point z_0* s'il existe un domaine $G \ni z_0$ dans lequel la fonction $f(z)$ est définie et analytique.

10.12. Exemples.

a. La fonction $f(z) = z$ est analytique partout dans le plan C_z puisque l'on a

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{z - z_0}{z - z_0} = 1,$$

et $f'(z) = 1$ pour tout z_0 .

b. La fonction $f(z) = \bar{z} = x - iy$ est dérivable en tout point z_0 par rapport à n'importe quelle demi-droite issue de z_0 , parce que

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{\bar{z} - \bar{z}_0}{z - z_0} = -2 \arg(z - z_0),$$

donc

$$f'_E(z_0) = -2 \arg(z - z_0).$$

Par contre, la même relation montre que $f(z) = \bar{z}$ n'est pas dérivable par rapport à un ensemble E contenant deux demi-droites différentes issues du point z_0 . A fortiori, $\bar{z}$ n'est une fonction analytique de z en aucun point $z_0 \in C_1$.

10.13. Propriétés générales de la dérivée. En raisonnant comme dans 7.13 et 7.15 on obtient les règles suivantes d'opérations sur les fonctions dérivables par rapport à un ensemble E en un point z_0 .

a. Si deux fonctions $f(z)$ et $g(z)$ sont dérivables pour $z = z_0$ par rapport à l'ensemble E , alors les fonctions $\alpha f(z)$ (α est un nombre complexe quelconque), $f(z) + g(z)$, $f(z)g(z)$, $\frac{f(z)}{g(z)}$ le sont aussi (dans le dernier cas, on suppose que $g(z_0) \neq 0$). Les formules suivantes sont alors valables (pour $z = z_0$):

$$(\alpha f(z))'_E = \alpha f'_E(z); \quad (1)$$

$$(f(z) + g(z))'_E = f'_E(z) + g'_E(z); \quad (2)$$

$$(f(z)g(z))'_E = f'_E(z)g(z) + f(z)g'_E(z); \quad (3)$$

$$\left(\frac{f(z)}{g(z)}\right)'_E = \frac{f'_E(z)g(z) - f(z)g'_E(z)}{g^2(z)}. \quad (4)$$

b. Si une fonction $w = f(z)$ est dérivable par rapport à l'ensemble E pour $z = z_0$ et une fonction $p = \varphi(w)$ est définie dans un ensemble F contenant tous les points $w = f(z)$ pour $z \in E$ assez proches de z_0 et dérivable par rapport à F pour $w = w_0 = f(z_0)$, alors la fonction composée $p(z) = \varphi(f(z))$ est dérivable par rapport à E pour $z = z_0$, et l'on a

$$p'_E(z_0) = \varphi'_F(w_0) \cdot f'_E(z_0).$$

c. Il résulte de (1), (3) et 10.12.a que *tout polynôme* $P(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n$ *est une fonction analytique ayant pour dérivée*

$$\begin{aligned} P'(z) &= (a_0 z^n + a_1 z^{n-1} + \dots + a_n)' = \\ &= n a_0 z^{n-1} + (n-1) a_1 z^{n-2} + \dots + a_{n-1}. \end{aligned}$$

d. Il résulte, ensuite, de (4) que *toute fonction rationnelle* c'est-à-dire tout rapport $\frac{P(z)}{Q(z)}$ de deux polynômes $P(z)$ et $Q(z)$, *est analytique partout sauf aux racines du dénominateur* $Q(z)$.

10.14. Dérivées d'ordres supérieurs. Les dérivées d'ordres supérieurs d'une fonction $f(z)$ par rapport à un ensemble E sont définies par récurrence, tout comme dans le cas d'une

variable réelle :

$$f_E^{(n+1)}(z) = [f_E^{(n)}(z)]'_E \quad (n=1, 2, \dots)$$

si les dérivées en question existent. Une fonction $f(z)$ qui a ses dérivées successives, y compris la n -ième, par rapport à un ensemble E est dite n fois *dérivable par rapport à l'ensemble E* .

Pour deux fonctions $f(z)$ et $g(z)$ n fois dérivables par rapport à un ensemble E , on a les formules :

$$[\alpha f(z)]_E^{(n)} = \alpha f_E^{(n)}(z) \quad (\alpha \text{ est un nombre complexe quelconque}) ; \quad (1)$$

$$[f(z) + g(z)]_E^{(n)} = f_E^{(n)}(z) + g_E^{(n)}(z) ; \quad (2)$$

$$[f(z) \cdot g(z)]_E^{(n)} = \sum_{k=0}^n \frac{n!}{k!(n-k)!} f_E^{(k)}(z) g_E^{(n-k)}(z). \quad (3)$$

Ces formules se démontrent de même façon que celles de 8.12 pour le cas d'une variable réelle.

10.15. Différentielle.

a. Comme dans le cas d'une variable réelle (7.31), l'accroissement d'une fonction $w = f(z)$ dérivable par rapport à un ensemble E en un point z_0 peut être mis sous la forme

$$f(z) - f(z_0) = f'_E(z_0)(z - z_0) + \varepsilon(z, z_0)(z - z_0),$$

où $\varepsilon(z, z_0)$ tend vers zéro pour $z \rightarrow z_0$. Le premier terme représente la *partie linéaire principale* de l'accroissement. La quantité $z - z_0$ est désignée par $d_E z$ ou par dz tout court, $f'_E(z_0) \cdot (z - z_0)$ par dw ou $dw(z_0)$. Donc, on peut noter la dérivée $f'_E(z_0)$ comme suit

$$f'_E(z_0) = \frac{dw}{d_E z} = \frac{dw}{dz}.$$

b. Tout comme dans 7.33-7.34, on établit les formules suivantes :

Quelles que soient une fonction $w = f(z)$ dérivable par rapport à un ensemble E et une constante complexe α , on a

$$d(\alpha w) = \alpha dw ; \quad (1)$$

quelles que soient les fonctions w_1 et w_2 dérivables par rapport à un ensemble E , on a

$$d(w_1 + w_2) = dw_1 + dw_2, \quad (2)$$

$$d(w_1 w_2) = w_1 dw_2 + dw_1 \cdot w_2 ; \quad (3)$$

$$d\left(\frac{w_1}{w_2}\right) = \frac{w_2 dw_1 - w_1 dw_2}{w_1^2} \quad (\text{pour } w_2(z_0) \neq 0). \quad (4)$$

c. Si $w = f(z)$ est dérivable par rapport à E pour $z = z_0$, et si une fonction $p = \varphi(w)$ est définie sur un ensemble F contenant tous

les points $w = f(z)$ et dérivable par rapport à F pour $w = w_0 = f(z_0)$, alors $dp = \varphi'(w) dw = \varphi'(w) w'(z) dz$, de sorte que la différentielle d'une fonction ne change pas de forme lorsque la variable indépendante devient fonction d'une autre variable indépendante.

10.16*. Conditions de Cauchy - Riemann. Donnons d'abord la définition d'une dérivée partielle.

Une fonction $F(x, y)$ étant définie dans un domaine G du plan des variables réelles x, y , le nombre

$$\lim_{h \rightarrow 0} \frac{F(x_0 + h, y_0) - F(x_0, y_0)}{h},$$

s'il existe, est appelé *dérivée partielle de la fonction $F(x, y)$ par rapport à la variable x au point $(x_0, y_0) \in G$* . D'une façon analogue, le nombre

$$\lim_{h \rightarrow 0} \frac{F(x_0, y_0 + h) - F(x_0, y_0)}{h}$$

est la *dérivée partielle de la fonction $F(x, y)$ par rapport à la variable y au point $(x_0, y_0) \in G$* . Le premier de ces nombres est désigné par $\frac{\partial F(x_0, y_0)}{\partial x}$, le deuxième par $\frac{\partial F(x_0, y_0)}{\partial y}$ ou, brièvement, par $F_x = F_x(x_0, y_0)$ et $F_y = F_y(x_0, y_0)$ respectivement.

Soit $w = f(z) = u(z) + iv(z) = u(x, y) + iv(x, y)$ une fonction dérivable pour $z = z_0 = x_0 + iy_0$ par rapport à un ensemble E qui contient le point z_0 ainsi que quelques segments contenant ce point et parallèles aux axes des coordonnées.

Posons $z = z_0 + h \in E$, h étant réel. Alors $z = z_0 + h = (x_0 + h, y_0)$ et l'on a

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} + i \frac{v(x_0 + h, y_0) - v(x_0, y_0)}{h}. \quad (1)$$

L'existence de la limite du rapport (1) implique l'existence des limites

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} &= \frac{\partial u(x_0, y_0)}{\partial x}, \\ \lim_{h \rightarrow 0} \frac{v(x_0 + h, y_0) - v(x_0, y_0)}{h} &= \frac{\partial v(x_0, y_0)}{\partial x} \end{aligned}$$

et l'égalité

$$f'_E(z_0) = \frac{\partial u(x_0, y_0)}{\partial x} + i \frac{\partial v(x_0, y_0)}{\partial x}. \quad (2)$$

Posons $z = z_0 + ih \in E$, h étant réel. Alors $z = z_0 + ih = (x_0, y_0 + h)$ et l'on a

$$\begin{aligned} \frac{f(z) - f(z_0)}{z - z_0} &= \frac{u(x_0, y_0 + h) - u(x_0, y_0)}{ih} + i \frac{v(x_0, y_0 + h) - v(x_0, y_0)}{ih} = \\ &= \frac{u(x_0, y_0 + h) - u(x_0, y_0)}{ih} + \frac{v(x_0, y_0 + h) - v(x_0, y_0)}{h}. \end{aligned}$$

L'existence de la dérivée $f'_E(z_0)$ implique l'existence des limites

$$\lim_{h \rightarrow 0} \frac{u(x_0, y_0 + h) - u(x_0, y_0)}{h} = \frac{\partial u(x_0, y_0)}{\partial y},$$

$$\lim_{h \rightarrow 0} \frac{v(x_0, y_0 + h) - v(x_0, y_0)}{h} = \frac{\partial v(x_0, y_0)}{\partial y}$$

et l'égalité

$$f'_E(z_0) = -i \frac{\partial u(x_0, y_0)}{\partial y} + \frac{\partial v(x_0, y_0)}{\partial y}. \quad (3)$$

En comparant (2) et (3) on voit que, en tout point z_0 où $f'_E(z_0)$ existe, on a les relations

$$\frac{\partial u(x_0, y_0)}{\partial x} = \frac{\partial v(x_0, y_0)}{\partial y},$$

$$\frac{\partial v(x_0, y_0)}{\partial x} = -\frac{\partial u(x_0, y_0)}{\partial y} \quad (4)$$

appelées *conditions de Cauchy-Riemann*.

10.17*. En particulier, si une fonction $f(z)$ est analytique dans un voisinage d'un point z_0 , alors les conditions de Cauchy-Riemann sont satisfaites en tout point de ce voisinage. Inversement, si les conditions de Cauchy-Riemann sont satisfaites dans un voisinage du point z_0 et si les fonctions $\frac{\partial u}{\partial x}$, $\frac{du}{dy}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$ sont continues, alors la fonction $f(z)$ est analytique dans ce voisinage. En effet, soit un $h = p + iq$ si petit que $z_0 + h$ appartient au voisinage mentionné. Formons la différence

$$\begin{aligned} f(z_0 + h) - f(z_0) &= \\ &= [u(x + p, y + q) + iv(x + p, y + q)] - [u(x, y) + iv(x, y)] = \\ &= [u(x + p, y + q) - u(x, y)] + i[v(x + p, y + q) - v(x, y)]. \end{aligned} \quad (1)$$

Transformons les expressions entre crochets. Nous avons

$$\begin{aligned} u(x + p, y + q) - u(x, y) &= \\ &= [u(x + p, y + q) - u(x, y + q)] + [u(x, y + q) - u(x, y)]. \end{aligned}$$

Appliquons la formule de Lagrange 7.44 à la première différence considérée comme fonction de x :

$$u(x + p, y + q) - u(x, y + q) = u_x(x + \theta_1 p, y + q) p,$$

où u_x est la dérivée par rapport à x et θ_1 un nombre entre 0 et 1. D'une manière analogue, en appliquant la formule de Lagrange à la deuxième différence, fonction de y , on a

$$u(x, y + q) - u(x, y) = u_y(x, y + \theta_2 q) q.$$

En effectuant les transformations analogues de l'autre crochet dans (1) on aboutit à la relation

$$\begin{aligned} f(z+h) - f(z) &= [u_x(x + \theta_1 p, y + q) p + \\ &+ u_y(x, y + \theta_2 q) q] + i [v_x(x + \theta_3 p, y + q) p + \\ &+ v_y(x, y + \theta_4 q) q]. \end{aligned}$$

En se servant des conditions de Cauchy-Riemann, on met cette différence sous la forme

$$f(z+h) - f(z) = [u_x(x + \theta_1 p, y + q) p + u_y(x, y + \theta_2 q) q] + i [u_x(x, y + \theta_4 q) q - u_y(x + \theta_3 p, y + q) p].$$

En vertu de l'hypothèse sur la continuité des fonctions u_x et u_y , on a ensuite :

$$f(z+h) - f(z) = u_x(x, y) (p + iq) - i u_y(x, y) (p + iq) + e_1(p, q) p + e_2(p, q) q,$$

où $e_1(p, q)$ et $e_2(p, q)$ tendent vers 0 avec p et q . Par conséquent,

$$\frac{f(z+h) - f(z)}{h} = u_x(x, y) - i u_y(x, y) + \frac{e_1(p, q) p + e_2(p, q) q}{p + iq}. \quad (2)$$

Estimons le dernier terme. On a $\left| \frac{p}{p+iq} \right| \leq 1$, $\left| \frac{q}{p+iq} \right| \leq 1$, d'où

$$\left| \frac{e_1(p, q) p + e_2(p, q) q}{p + iq} \right| \leq |e_1(p, q)| + |e_2(p, q)|.$$

La dernière quantité tendant vers 0 avec p et q , l'expression (2) a $u_x - i u_y$ pour limite lorsque $h \rightarrow 0$. Ceci veut dire que la dérivée de la fonction $f(z)$, pour $z = z_0$, existe et vaut $u_x(z_0) - i u_y(z_0)$. Le théorème est démontré.

Si l'on introduit la notion de fonction différentiable de deux variables, on peut formuler les conditions de Cauchy-Riemann sous forme de critère nécessaire et suffisant de l'analyticité. (Cf. exercice 6.)

10.18*. Dans la suite (10.34), on verra que toute fonction analytique $f(z)$ possède ses dérivées d'ordre supérieur quelconque. Quelles sont les conditions imposées aux parties réelle et imaginaire de $f(z)$ par l'existence de la dérivée seconde? Désignons par $\frac{\partial^2 u(x, y)}{\partial x^2}$ la dérivée partielle seconde de la fonction $u(x, y)$ par rapport à la variable x (c'est-à-dire la dérivée seconde de la fonction $u(x, y)$ en tant que fonction de x seul, avec un y constant). La dérivée partielle seconde $\frac{\partial^2 u(x, y)}{\partial y^2}$ a un sens analogue.

Les mêmes raisonnements que dans 10.16 nous conduisent aux relations suivantes

$$\begin{aligned} f''(z) &\equiv (f'(z))' = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) = \frac{\partial^2 u}{\partial x^2} + i \frac{\partial^2 v}{\partial x^2} = \\ &= \frac{1}{i} \frac{\partial}{\partial y} \left(\frac{1}{i} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y} \right) = -\frac{\partial^2 u}{\partial y^2} - i \frac{\partial^2 v}{\partial y^2}. \end{aligned}$$

En égalant les parties réelles et imaginaires on trouve

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

L'équation

$$\frac{\partial^2 \varphi(x, y)}{\partial x^2} + \frac{\partial^2 \varphi(x, y)}{\partial y^2} = 0$$

s'appelle *équation de Laplace*; une fonction réelle continue qui lui satisfait dans un domaine G s'appelle *fonction harmonique* dans G . Nous voyons que les parties réelle et imaginaire d'une fonction analytique dans un domaine G sont des fonctions harmoniques dans G .

Inversement, toute fonction $u(x, y)$ harmonique dans un domaine G est la partie réelle d'une fonction analytique $f(z)$ (et, bien entendu, la partie imaginaire de la fonction analytique $if(z)$). On le verra dans la Troisième partie (14.49.c).

Deux fonctions harmoniques $u(x, y)$ et $v(x, y)$ représentant respectivement les parties réelle et imaginaire d'une fonction analytique $f(z)$, c'est-à-dire liées par les conditions de Cauchy-Riemann, sont appelées *fonctions harmoniques conjuguées*.

§ 10.2. Intégrales curvilignes de fonctions complexes

10.21. Nous considérons ici les intégrales curvilignes prises le long des chemins lisses par morceaux (9.91) dans le domaine complexe.

Dans 9.91, on a défini l'intégrale curviligne de Stieltjes

$$\int_L f(M) dg(M),$$

où $f[M(t)]$ est une fonction continue par morceaux, $g[M(t)]$ une fonction continûment dérivable, les deux étant réelles. Cette intégrale se réduit à l'intégrale définie usuelle

$$\int_a^b f[M(t)] dg[M(t)] = \int_a^b f[M(t)] \frac{dg[M(t)]}{dt} dt,$$

où $[a, b]$ est l'intervalle de variation du paramètre t qui correspond au chemin L . Nous allons étendre la définition au cas où les fonctions $f(M)$ et $g(M)$ prennent des valeurs complexes.

Si la fonction $f(M)$ prend les valeurs complexes $f(M) = f_1(M) + if_2(M)$, où $f_1(M)$ et $f_2(M)$ sont réelles, alors l'intégrale curviligne de Stieltjes est définie par la formule

$$\int_L f(M) dg(M) = \int_L f_1(M) dg(M) + i \int_L f_2(M) dg(M)$$

(la fonction $g(M)$ est réelle). D'une façon analogue, si $f(M)$ est réelle et $g(M)$ complexe, de sorte que $g(M) = g_1(M) + ig_2(M)$, alors on pose

$$\int_L f(M) dg(M) = \int_L f(M) dg_1(M) + i \int_L f(M) dg_2(M).$$

Lorsque les deux fonctions $f(M) = f_1(M) + if_2(M)$ et $g(M) = g_1(M) + ig_2(M)$ prennent des valeurs complexes, on est amené, naturellement, à la définition

$$\begin{aligned} \int_L f(M) dg(M) = & \int_L f_1(M) dg_1(M) - f_2(M) dg_2(M) + \\ & + i \int_L f_1(M) dg_2(M) + f_2(M) dg_1(M). \end{aligned}$$

Dans le cas du plan ($n = 2$), pour $g(x, y) = z = x + iy$ et $f(x, y) = u(x, y) + iv(x, y)$, on obtient la formule

$$\int_L f(x, y) dz = \int_L u dx - v dy + i \int_L u dy + v dx. \quad (1)$$

Dans ce cas, il est possible de donner une définition directe de l'intégrale complexe (1), notamment

$$\int_L f(x, y) dz = \lim_{d(\Pi) \rightarrow 0} \sum_{j=0}^{m-1} f(x_j, y_j) \Delta z_j, \quad (2)$$

où Π est une partition du chemin L , $z_j = (x_j, y_j)$ le j -ième point de division et $\Delta z_j = z_{j+1} - z_j$. En effet, on a

$$\begin{aligned} \sum_{j=0}^{m-1} f(x_j, y_j) \Delta z_j = & \sum_{j=0}^{m-1} u(x_j, y_j) \Delta x_j - \sum_{j=0}^{m-1} v(x_j, y_j) \Delta y_j + \\ & + i \sum_{j=0}^{m-1} u(x_j, y_j) \Delta y_j + i \sum_{j=0}^{m-1} v(x_j, y_j) \Delta x_j, \end{aligned}$$

et cette expression tend exactement vers (1) lorsque $d(\Pi) \rightarrow 0$. De la représentation (2) on déduit aisément les propriétés principales de l'intégrale: la relation

$$\int_L [\alpha_1 f_1(x, y) + \alpha_2 f_2(x, y)] dz = \alpha_1 \int_L f_1(x, y) dz + \alpha_2 \int_L f_2(x, y) dz \quad (3)$$

quelles que soient deux fonctions intégrables $f_1(x, y)$ et $f_2(x, y)$ et deux constantes complexes α_1 et α_2 ; la relation

$$\int_L f(x, y) dz = \int_{L_1} f(x, y) dz + \int_{L_2} f(x, y) dz, \quad (4)$$

où L est un chemin lisse par morceaux, formé de deux parties adjacentes L_1 et L_2 avec un même sens de parcours; la majoration

$$\left| \int_L f(x, y) dz \right| \leq \int_L |f(x, y)| ds \leq \sup_{z \in L} |f(x, y)| \cdot s(L), \quad (5)$$

où $s(L)$ est la longueur du chemin L .

Parfois, pour préciser une propriété du chemin d'intégration, on emploie des symboles spéciaux

$$\oint_L, \oint_L, \oint_L, \int_L, \int_L, \int_L, \text{ etc.}$$

Le premier symbole signifie que l'intégration se fait le long d'un contour fermé; les deuxième et troisième précisent le sens de parcours; les quatrième et cinquième montrent que l'intégration s'effectue le long d'une droite et indiquent le sens de parcours, le sixième que le contour d'intégration représente une demi-circonférence parcourue dans le sens indiqué. En général, le sens d'un symbole est révélé par le texte ou par une figure.

10.22. Considérons une suite de fonctions $f_n(z)$ définies sur un chemin lisse par morceaux $L \subset C_z$. Supposons que les fonctions $f_n(z)$ y convergent uniformément vers une fonction $f(z)$. Si les fonctions $f_n(z)$ sont continues sur L , alors il en est de même de la fonction limite $f(z)$ (5.96).

a. Dans l'hypothèse formulée ci-dessus, on a

$$\int_L f(z) dz = \lim \int_L f_n(z) dz.$$

En effet, il résulte de 10.21 (5) que

$$\left| \int_L [f(z) - f_n(z)] dz \right| \leq \max_L |f(z) - f_n(z)| \cdot s(L),$$

ce qui conduit au résultat cherché.

b. Si, de plus, les fonctions $f_n(z)$ et $f(z)$ dépendent d'un paramètre λ qui parcourt un ensemble Λ , de sorte que l'on puisse noter $f_n(z) \equiv f_n(z, \lambda)$, $f(z) \equiv f(z, \lambda)$, et si la convergence de $f_n(z, \lambda)$ vers $f(z, \lambda)$ est uniforme par rapport à λ , c'est-à-dire que

$$\sup_{z, \lambda} |f_n(z, \lambda) - f(z, \lambda)| \rightarrow 0,$$

alors les fonctions

$$F_n(\lambda) = \int_L f_n(z, \lambda) dz$$

convergent elles aussi vers la fonction

$$F(\lambda) = \int_L f(z, \lambda) dz$$

uniformément par rapport à $\lambda \in \Lambda$.

C'est encore une conséquence de l'estimation 10.21 (5):

$$\left| \int_L [f(z, \lambda) - f_n(z, \lambda)] dz \right| \leq \sup_{z, \lambda} |f_n(z, \lambda) - f(z, \lambda)| \cdot s(L).$$

10.23. Théorème de la moyenne. Si $f(z)$ est continue par morceaux sur un chemin lisse $L = \{z: z = z(t), a \leq t \leq b\}$, et si $z'(t_0) \neq 0$, la fonction $f(z)$ étant continue en $z = z_0 = z(t_0)$, alors

$$\lim_{z \rightarrow z_0} \frac{1}{z - z_0} \int_{L(z, z_0)} f(z) dz = f(z_0),$$

où $L(z, z_0)$ est la partie du chemin L entre les points z et z_0 .

Démonstration. En cas de $f(z) \equiv f(z_0) = \text{const}$, la conclusion du théorème résulte immédiatement de la définition de l'intégrale. On peut donc se borner au cas où $f(z_0) = 0$ (sinon, on retranche la constante $f(z_0)$ de la fonction $f(z)$). Pour un $\varepsilon > 0$ donné, on trouve un $\delta > 0$ tel que $z \in L$, $|z - z_0| < \delta$ implique $|f(z)| < \varepsilon$. Alors, d'après 10.21 (5), on a

$$\left| \int_{L(z, z_0)} f(z) dz \right| \leq \varepsilon \cdot s(L(z, z_0)).$$

Donc

$$\frac{\left| \int_{L(z, z_0)} f(z) dz \right|}{|z - z_0|} \leq \varepsilon \frac{s(L(z, z_0))}{|z - z_0|}. \quad (1)$$

Or, le chemin L étant lisse, le rapport de la longueur d'une partie du chemin à celle de la corde qui la sous-tend converge vers 1, pour $z \rightarrow z_0$ (9.63.f). Par conséquent, pour $|z - z_0|$ assez petits, la majoration (1) peut être remplacée par

$$\frac{1}{|z - z_0|} \left| \int_{L(z, z_0)} f(z) dz \right| \leq 2\varepsilon,$$

d'où le théorème.

10.24. Soit $f(z, \zeta)$ une fonction de deux variables complexes; la variable $z = x + iy$ parcourt un chemin lisse par morceaux L dans le plan des z et la variable $\zeta = \xi + i\eta$ un chemin lisse par morceaux Λ dans le plan des ζ . On suppose que la fonction $f(z, \zeta)$ soit continue par rapport à l'ensemble des variables, c'est-à-dire que, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que les relations $|z - z'| \leq \delta$, $|\zeta - \zeta'| \leq \delta$, $z, z' \in L$, $\zeta, \zeta' \in \Lambda$ impliquent $|f(z', \zeta') - f(z, \zeta)| < \varepsilon$. Alors chacune des intégrales

$$I(\zeta) = \int_L f(z, \zeta) dz, \quad J(z) = \int_\Lambda f(z, \zeta) d\zeta \quad (1)$$

représente une fonction continue de la variable correspondante, et l'on a

$$\int_L \left\{ \int_\Lambda f(z, \zeta) d\zeta \right\} dz = \int_\Lambda \left\{ \int_L f(z, \zeta) dz \right\} d\zeta. \quad (2)$$

Dans 9.82, on a démontré le théorème analogue pour les fonctions réelles. Or, on a vu plus haut qu'une intégrale prise le long d'un chemin complexe est une combinaison linéaire d'intégrales par rapport à une variable réelle. Notre affirmation résulte donc directement du théorème 9.82.

10.25.a. Soit $w = f(z)$ une fonction dérivable par rapport à un ensemble E contenant un chemin lisse par morceaux L dont l'origine est z_0 et l'extrémité z_1 . Alors, la formule de Newton-Leibniz

$$\int_L f'_L(z) dz = f(z_1) - f(z_0)$$

est valable.

En effet, soit $z = z(t)$, $t_0 \leq t \leq t_1$, une représentation paramétrique du chemin L , avec la fonction $z(t)$ lisse par morceaux, et soit $F(t) = f(z(t))$. Alors, d'après le théorème sur la dérivée d'une fonction composée 10.13.b, on a

$$f'_L(z) dz(t) = dF(t).$$

d'où

$$\int_L f'_L(z) dz = \int_{t_0}^{t_1} dF(t) = F(t_1) - F(t_0) = f(z_1) - f(z_0),$$

ce qu'il fallait démontrer.

b. En particulier, si $w = f(z)$ est une fonction analytique dans un domaine G , on a

$$\int_L f'(z) dz = f(z_1) - f(z_0)$$

pour tout chemin lisse par morceaux L qui joint les points z_1 et z_0 (et qui est inclus entièrement dans G). Il en résulte que l'intégrale de $f'(z)$ prise le long de tout contour fermé lisse par morceaux $L \subset G$ vaut 0.

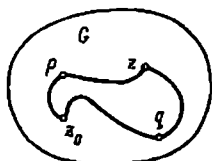
c. Ainsi, l'intégrale le long d'un contour fermé L d'un polynôme

$$P(z) = a_0 + a_1 z + \dots + a_n z^n$$

est nulle, puisque le polynôme $P(z)$ est la dérivée du polynôme

$$P_1(z) = a_0 z + a_1 \frac{z^2}{2} + \dots + a_n \frac{z^{n+1}}{n+1}.$$

10.26. La propriété 10.25.b est inversible. Soit $\varphi(z)$ une fonction donnée dans un domaine connexe (10.11.e) G et telle que son intégrale le long de tout chemin fermé lisse par morceaux L dans G est nulle :



$$\oint_L \varphi(z) dz = 0.$$

Il existe une fonction $f(z)$ analytique dans le domaine G et telle que $f'(z) = \varphi(z)$.

Pour le démontrer, fixons dans le domaine G un point quelconque z_0 et considérons un autre point arbitraire z , ainsi qu'un chemin lisse par morceaux $L_z \subset G$ joignant z_0 et z . Posons

$$f(z) = \int_{L_z} \varphi(\zeta) d\zeta.$$

La valeur $f(z)$ ne dépend pas du choix du chemin menant de z_0 à z . En effet, s'il y a deux chemins $z_0 p z$ et $z_0 q z$, on peut en former le contour fermé $z q z_0 p z$ (fig. 10.1), l'intégrale le long de ce contour étant nulle par hypothèse; par conséquent, on a

$$\int_{z_0 p z} \varphi(\zeta) d\zeta = \int_{z_0 q z} \varphi(\zeta) d\zeta - \int_{z q z_0 p z} \varphi(\zeta) d\zeta = \int_{z q z} \varphi(\zeta) d\zeta$$

Démontrons que $f(z)$ est une fonction analytique.

Pour un z donné, choisissons un nombre $\rho > 0$ tel que le cercle de rayon ρ centré en z appartienne au domaine G . Soit, ensuite, un nombre complexe h , $|h| < \rho$. Puisque le chemin qui va de z_0 à $z + h$ peut être choisi arbitrairement, nous le construisons en ajoutant au chemin déjà fixe de z_0 à z le segment de la droite passant par z et $z + h$. Nous avons alors

$$f(z+h) - f(z) = \int_{z_0}^{z+h} \varphi(\zeta) d\zeta - \int_{z_0}^z \varphi(\zeta) d\zeta = \int_z^{z+h} \varphi(\zeta) d\zeta$$

et

$$\frac{f(z+h) - f(z)}{h} \rightarrow \frac{1}{h} \int_z^{z+h} \varphi(\zeta) d\zeta.$$

Cette quantité tend vers $\varphi(z)$, lorsque $h \rightarrow 0$, en vertu de 10.23, d'où l'analyticité de $f(z)$.

10.27. Soit $f(z, \zeta)$ une fonction de deux variables complexes: la variable $z = x + iy$ parcourt un domaine G dans le plan des z , la variable ζ un chemin liasse par morceaux Λ . On suppose que la fonction $f(z, \zeta)$ soit analytique dans G pour tout $\zeta \in \Lambda$ et que la fonction $f(z, \zeta)$ ainsi que sa dérivée $f'_z(z, \zeta)$ soient continues par rapport à l'ensemble des variables $z \in G$, $\zeta \in \Lambda$. Formons la fonction

$$\Phi(z) = \int_{\Lambda} f(z, \zeta) d\zeta.$$

T h é o r è m e. Dans les conditions formulées ci-dessus, la fonction $\Phi(z)$ est analytique dans le domaine G .

D é m o n s t r a t i o n. Considérons la fonction

$$\varphi(z) = \int_{\Lambda} f'_z(z, \zeta) d\zeta.$$

Nous disons que, pour tout contour fermé $L \subset G$, on a

$$\int_L \varphi(z) dz = 0.$$

En effet, d'après 10.24 et 10.25.b, on a

$$\int_L \varphi(z) dz = \int_L \left\{ \int_{\Lambda} f'_z(z, \zeta) d\zeta \right\} dz = \int_{\Lambda} \left\{ \int_L f'_z(z, \zeta) dz \right\} d\zeta = 0.$$

En vertu de 10.26, il existe une fonction $\Phi_0(z)$ analytique dans le domaine G dont $\varphi(z)$ est la dérivée, et, de plus, toujours d'après 10.24, on a

$$\begin{aligned} \Phi_0(z) &= \int_{z_0}^z \varphi(\alpha) d\alpha = \int_{z_0}^z \left\{ \int_{\Lambda} f'_z(\alpha, \zeta) d\zeta \right\} d\alpha = \\ &= \int_{\Lambda} \left\{ \int_{z_0}^z f'_z(\alpha, \zeta) d\alpha \right\} d\zeta = \int_{\Lambda} f(z, \zeta) d\zeta - \int_{\Lambda} f(z_0, \zeta) d\zeta. \end{aligned}$$

La différence de $\Phi_0(z)$ et $\Phi(z)$ étant constante, nous voyons que $\Phi(z)$, de même que $\Phi_0(z)$, est une fonction analytique dans le domaine G .

10.28. Considérons une fonction de la forme

$$F(z) = \int_{\Lambda} \frac{f(\zeta)}{\zeta - z} d\zeta, \quad (1)$$

où $f(\zeta)$ est continue sur un chemin lisse par morceaux Λ et le point z est extérieur au chemin. Ainsi, la fonction $F(z)$ est définie dans le domaine G qui est le complémentaire du contour Λ par rapport à tout le plan G_z . Désignons par G_ρ l'ensemble des points du domaine G pour lesquels $\sup |z - \zeta| > \rho$. Quel que soit un point $z \in G$, il appartient à un G_ρ si ρ est assez petit. Notons maintenant que la fonction

$$f(\zeta, z) = \frac{f(\zeta)}{\zeta - z}$$

est analytique en z dans le domaine G et, qui plus est, possède dans G ses dérivées de tous les ordres :

$$f^{(n)}_z(z, \zeta) = \frac{n! f(\zeta)}{(\zeta - z)^{n+1}}.$$

Il est évident que ces dérivées sont continues par rapport à l'ensemble des variables $\zeta \in \Lambda$, $z \in G_\rho$, pour tout ρ , et analytiques en z dans le domaine G . En appliquant le théorème 10.27 on obtient :

T h é o r è m e. *La fonction $F(z)$ définie par l'expression (1) est analytique dans le domaine G et, qui plus est, y possède ses dérivées de tous les ordres qui sont analytiques elles aussi. On a*

$$F^{(n)}(z) = n! \int_{\Lambda} \frac{f(\zeta) d\zeta}{(\zeta - z)^{n+1}}$$

Une intégrale de la forme (1) est appelée *intégrale du type Cauchy*.

§ 10.3. Théorème de Cauchy et ses conséquences

On réunit, dans le présent paragraphe, les résultats fondamentaux de la théorie des fonctions analytiques (qui sont trouvés, pour la plupart, par Cauchy).

10.31. Un domaine connexe G (10.11.e) est dit *simplement connexe* si tout polygone dont le contour est inclus dans G peut être mis sous la forme d'une réunion finie de triangles qui se trouvent dans G avec leur intérieur. Par exemple, un cercle ouvert, ainsi que tout domaine convexe (qui contient, avec ses deux points quelconques, le segment qui les joint) sont simplement connexes, tandis qu'une couronne circulaire $r_1 < |z| < r_2$ ne l'est pas (fig. 10.2).

10.32. Théorème de Cauchy. *Pour une fonction $f(z)$ analytique dans un domaine simplement connexe V , on a*

$$\oint_L f(z) dz = 0,$$

où L est n'importe quel contour fermé contenu entièrement dans le domaine V .

Démonstration. Nous avons vu dans 9.95 que l'intégrale prise le long d'une courbe est la limite des intégrales prises

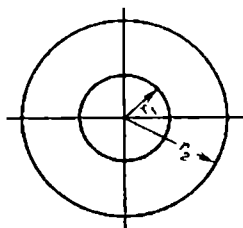


Fig. 10.2.

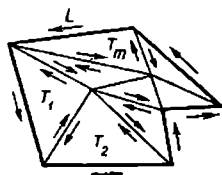


Fig. 10.3.

le long de lignes polygonales; il suffit donc de démontrer le théorème de Cauchy pour un polygone. Le domaine V étant simplement connexe, on peut partager le polygone L en triangles $T_1, \dots, T_m$ sans sortir du domaine; on aura (fig. 10.3)

$$\oint_L f(z) dz = \oint_{T_1} f(z) dz + \dots + \oint_{T_m} f(z) dz$$

(les intégrales sur les segments à l'intérieur du polygone se réduisent, ces segments étant parcourus deux fois dans les sens inverses); il suffit donc de démontrer le théorème de Cauchy pour un triangle.

Considérons un triangle quelconque T dans le domaine V et désignons par ρ son périmètre, par c la valeur absolue de l'intégrale

$$\oint f(z) dz.$$

Partageons le triangle T en quatre triangles égaux T_1, T_2, T_3, T_4 en joignant les milieux de ses côtés par les segments de droite (fig. 10.4). On aura alors

$$\oint_T f(z) dz = \oint_{T_1} f(z) dz + \oint_{T_2} f(z) dz + \oint_{T_3} f(z) dz + \oint_{T_4} f(z) dz.$$

Donc, pour l'un au moins des triangles, par exemple pour T_1 , on a l'inégalité

$$\left| \oint_{T_1} f(z) dz \right| \geq \frac{c}{4}.$$

Partageons le triangle T_1 encore une fois en quatre triangles égaux $T_{11}, \dots, T_{14}$ et choisissons-en un, par exemple T_{11} , pour lequel

$$\left| \oint_{T_{11}} f(z) dz \right| \geq \frac{1}{4} \left| \oint_{T_1} f(z) dz \right| \geq \frac{c}{4^2}.$$

Si l'on continue indéfiniment ce processus, on aboutit à la suite des triangles emboîtés $T_1 = T_1, T_2 = T_{11}, \dots$ tels que

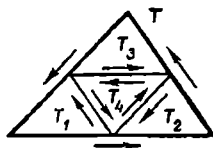


Fig. 10.4.

$$\left| \oint_{T_n} f(z) dz \right| \geq \frac{c}{4^n}. \quad (1)$$

Chaque triangle T_n (considéré avec son intérieur) est un ensemble compact dans le plan $C_z = R_2$ (5.14.e); en vertu de 3.98, il existe dans le plan un point z_0 commun à tous ces triangles.

La fonction $f(z)$ est dérivable en ce point comme partout dans le domaine V ; la définition de la dérivée nous donne que, quel que soit un $\varepsilon > 0$, il existe un $\delta > 0$ tel que, si $|z - z_0| < \delta$, alors

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + R_1(z, z_0), \quad (2)$$

où $|R_1(z, z_0)| < \varepsilon |z - z_0|$ (10.15). Le reste R_1 est une fonction au moins continue (en z), puisque les autres termes de l'égalité (2) le sont. Estimons maintenant l'intégrale de $f(z)$ le long du triangle T_n contenu dans le cercle de rayon δ centré en z_0 . En nous servant de la relation (2), nous avons

$$\oint_{T_n} f(z) dz = \oint_{T_n} f(z_0) dz + f'(z_0) \oint_{T_n} (z - z_0) dz + \oint_{T_n} R_1(z, z_0) dz.$$

Les deux premiers termes dans le second membre sont nuls d'après 10.25.c. Le périmètre du triangle T_n valant $\frac{\rho}{2^n}$, on a en appliquant l'estimation (2) du reste R_1 :

$$\left| \oint_{T_n} f(z) dz \right| = \left| \oint_{T_n} R_1(z, z_0) dz \right| \leq \varepsilon \frac{\rho}{2^n} \frac{\rho}{2^n}.$$

En comparant cette inégalité et (1) on voit que

$$\frac{c}{4^n} \leq e \frac{\rho^2}{4^n}, \quad c \leq e \rho^2.$$

Le nombre e étant arbitrairement petit, on a $c = 0$, et le théorème de Cauchy est démontré.

10.33. Formule de Cauchy.

a. Un cercle $|z - a| < R$ dépourvu du point $z = a$ fournit un exemple de domaine non simplement connexe. La fonction $\frac{1}{(z-a)^m}$ ($m \geq 1$) est analytique dans ce domaine (et ne l'est pas dans tout le cercle $|z - a| < R$). Calculons l'intégrale

$$\oint_L \frac{dz}{(z-a)^m},$$

où L est la circonférence de rayon $r < R$ centrée en a . Choisissons pour paramètre l'angle polaire φ ; alors $z - a = re^{i\varphi}$, $(z - a)^m = r^m e^{im\varphi}$, $dz = ire^{i\varphi} d\varphi$, et l'on a :

$$\oint_L \frac{dz}{(z-a)^m} = ir^{1-m} \int_{\varphi=0}^{2\pi} e^{i(1-m)\varphi} d\varphi =$$

$$= \begin{cases} ir^{1-m} \frac{e^{i(1-m)\varphi}}{i(1-m)} \Big|_0^{2\pi} = 0 & (m \neq 1), \end{cases} \quad (1)$$

$$= \begin{cases} i \int_0^{2\pi} d\varphi = 2\pi i & (m = 1). \end{cases} \quad (2)$$

Pour $m = 1$, l'intégrale vaut $2\pi i \neq 0$, ce qui montre, entre autres, que l'hypothèse de connexité simple dans le théorème de Cauchy est essentielle.

b. Signalons encore une propriété importante des intégrales suivant les contours fermés dans un domaine non simplement connexe. Si, dans un domaine V (que l'on ne suppose pas simplement connexe), une fonction analytique $f(z)$ est définie et si deux contours fermés lisses par morceaux L_1 et L_2 sont choisis d'une manière indiquée fig. 10.5 (ces contours et une ligne T lisse par morceaux qui les joint forment la frontière d'un domaine simplement connexe), alors on a

$$\oint_{L_1} f(z) dz = \oint_{L_2} f(z) dz. \quad (3)$$

En effet, ayant donné le domaine simplement connexe mentionné, établissons le sens de parcours de sa frontière comme suit : le contour

L_1 est parcouru dans le sens positif (contre l'aiguille d'une montre), le contour L_2 dans le sens négatif (celui de l'aiguille d'une montre), et la ligne T qui les joint est parcourue deux fois dans les sens inverses (fig. 10.5). D'après le théorème de Cauchy 10.32, on a

$$0 = \oint_L f(z) dz = \oint_{L_1} f(z) dz + \int_T f(z) dz + \int_{\bar{T}} f(z) dz + \oint_{L_2} f(z) dz,$$

d'où (3) si l'on se rappelle qu'une intégrale curviligne change de signe lorsqu'on inverse le sens de parcours du chemin d'intégration (9.91).

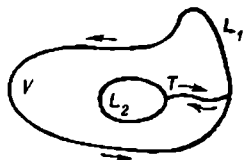


Fig. 10.5.

Le cas de la fig. 10.5 n'est pas le plus général. Tout de même, il est suffisant pour nos besoins ultérieurs.

e. Appliquons la propriété ci-dessus au calcul de l'intégrale

$$I = \oint_L \frac{f(z) dz}{z-a}$$

le long d'une courbe fermée L contenant le point a dans le sens positif. La fonction $f(z)$ est supposée analytique dans un domaine simplement connexe V contenant le point a et le contour L . La fonction $\frac{f(z)}{z-a}$ a, dans le cas général, une singularité pour $z = a$ (c.-à-d. peut ne pas y avoir de dérivée), donc l'intégrale I peut être non nulle. Néanmoins, la fonction $\frac{f(z)}{z-a}$ étant analytique à l'extérieur de a de même que $f(z)$, il est possible, d'après ce qu'on a vu plus haut, de remplacer le contour L par une circonférence d'un petit rayon ε centrée en a , sans changer la valeur de l'intégrale I . Posons

$$f(z) = f(a) + R(z).$$

La fonction $\frac{R(z)}{z-a}$ est continue, avec $f(z)$, pour $z \neq a$; elle tend vers $f'(a)$ lorsque $z \rightarrow a$; elle est donc bornée dans un voisinage du point a ; soit, par exemple, $\left| \frac{R(z)}{z-a} \right| \leq M$. On a alors

$$\oint_L \frac{f(z) dz}{z-a} = \oint_{|z-a|=\varepsilon} \frac{f(z) dz}{z-a} = f(a) \oint_{|z-a|=\varepsilon} \frac{dz}{z-a} + \oint_{|z-a|=\varepsilon} \frac{R(z)}{z-a} dz. \quad (4)$$

Le premier terme dans le dernier membre vaut $2\pi i \cdot f(a)$ d'après 10.33 (2). Pour le deuxième, on obtient en vertu de 10.21 (5):

$$\left| \oint_{|z-a|=\varepsilon} \frac{R(z)}{z-a} dz \right| \leq M \cdot 2\pi\varepsilon.$$

Ainsi, la deuxième terme tend vers 0 lorsque $\varepsilon \rightarrow 0$. Comme deux autres membres de (4) ne dépendent pas de ε , on obtient à la limite :

$$I = \oint_L \frac{f(z) dz}{z-a} = 2\pi i f(a)$$

ou bien

$$f(a) = \frac{1}{2\pi i} \oint_L \frac{f(z) dz}{z-a}. \quad (5)$$

La formule de Cauchy (5) est l'une des plus importantes formules de la théorie des fonctions analytiques : elle représente les valeurs d'une fonction analytique à l'intérieur d'un contour fermé L en fonction de ses valeurs sur ce contour.

10.34. Théorème. Une fonction $w = f(z)$ analytique dans un domaine G possède, dans ce domaine, ses dérivées de tous les ordres qui sont analytiques elles aussi. De plus, on a

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_L \frac{f(\zeta) d\zeta}{(\zeta-z)^{n+1}}. \quad (1)$$

où l'intégrale est prise le long de n'importe quelle courbe fermée L qui contourne le point z dans le sens positif et est contenue dans le domaine G .

En effet, la formule de Cauchy 10.33 (5) montre que $f(z)$ est représentable, dans G , par une intégrale du type Cauchy ; il reste d'appliquer le théorème 10.28.

10.35.a. Théorème. Toute fonction $w = f(z)$ analytique à l'intérieur d'un cercle $V = \{z : |z - z_0| \leq r\}$ se développe en une série de puissances

$$f(z) = \sum_0^{\infty} a_n (z - z_0)^n \quad (1)$$

qui converge à l'intérieur du cercle V .

Démonstration. Nous partons de la formule de Cauchy

$$f(z) = \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta) d\zeta}{\zeta - z}$$

qui est valable pour tout contour $\Lambda = \{\zeta : |\zeta - z_0| = r - \delta\}$ et pour tout point z intérieur à Λ . Dans ces conditions, on a :

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0} \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} = \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0} \right)^n.$$

La série obtenue converge uniformément sur Λ par rapport à ζ d'après l'estimation

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = \frac{|z - z_0|}{r - \delta} < 1$$

et le critère de Weierstrass 6.53. Vu le théorème 10.22.b, on a

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta)}{\zeta - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0} \right)^n d\zeta = \\ &= \sum_{n=0}^{\infty} (z - z_0)^n \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta) d\zeta}{(\zeta - z_0)^{n+1}}, \end{aligned}$$

ce qui fournit le développement cherché ainsi que l'expression du coefficient a_n :

$$a_n = \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta) d\zeta}{(\zeta - z_0)^{n+1}}. \quad (2)$$

b. En vertu de 10.34, nous avons

$$a_n = \frac{f^{(n)}(z_0)}{n!}. \quad (3)$$

Ceci permet de renforcer le théorème a:

T h é o r è m e. *Toute fonction $w = f(z)$ analytique à l'intérieur d'un cercle $V = \{z : |z - z_0| < r\}$ se développe, dans ce cercle, en la série de puissances*

$$f(z) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(z_0) (z - z_0)^n. \quad (4)$$

La série (4) s'appelle *série de Taylor pour la fonction analytique $f(z)$ dans le cercle V* .

e. **C o n s é q u e n c e.** *Si une fonction $f(z)$ est analytique dans un cercle $V = \{z : |z - z_0| < r\}$ et que $f^{(n)}(z_0) = 0$ ($n = 0, 1, \dots$), alors $f(z) \equiv 0$ dans le cercle V .*

10.36.a. Nous disons qu'une suite (série) de fonctions $f_n(z)$ définies dans un domaine G converge uniformément à l'intérieur de G si la convergence de cette suite (série) est uniforme sur tout compact Q inclus dans G .

Notons que le compact Q inclus dans G peut être contourné par une courbe fermée lisse par morceaux qui soit incluse elle aussi dans G . En effet, tout point $z_0 \in Q$ peut être recouvert par un cercle ouvert $V = \{z : |z - z_0| < \rho\}$ contenu, avec sa frontière, dans G . En vertu de 3.97, on peut dégager du recouvrement obtenu du

compact Q un recouvrement fini $V_1, \dots, V_m$. La frontière du domaine $V = V_1 \cup \dots \cup V_m$ représente un contour fermé lisse par morceaux qui est formé d'un nombre fini d'arcs de circonférence; par construction, il est inclus dans le domaine $G - Q$ et contourne Q , ce qu'il fallait démontrer.

b. Théorème. Si une suite $f_1(z), f_2(z), \dots$ de fonctions analytiques dans un domaine G converge uniformément à l'intérieur de G vers une fonction $f(z)$, alors $f(z)$ est aussi analytique dans G . La suite $f_1^{(n)}(z), f_2^{(n)}(z), \dots$ des dérivées d'un ordre quelconque converge uniformément à l'intérieur de G vers la dérivée correspondante $f^{(n)}(z)$ de la fonction $f(z)$.

Démonstration. On a, d'après la formule de Cauchy 10.33,

$$f_n(z) = \frac{1}{2\pi i} \oint_{\Lambda} \frac{f_n(\zeta) d\zeta}{\zeta - z}$$

pour toute courbe fermée $\Lambda \subset G$ contournant un point donné $z \in G$. En passant à la limite pour $n \rightarrow \infty$ et en se servant du théorème 10.22.a on trouve

$$f(z) = \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta) d\zeta}{\zeta - z}.$$

La fonction $f(z)$ est donc représentable par une intégrale du type Cauchy et, par conséquent, analytique dans G .

Pour démontrer la deuxième partie du théorème, considérons les représentations 10.34 pour les fonctions $f^{(m)}(z)$ et $f_n^{(m)}(z)$:

$$f^{(m)}(z) = \frac{n!}{2\pi i} \oint_L \frac{f(\zeta) d\zeta}{(\zeta - z)^{m+1}}, \quad f_n^{(m)}(z) = \frac{n!}{2\pi i} \oint_L \frac{f_n(\zeta) d\zeta}{(\zeta - z)^{m+1}},$$

où les intégrales sont prises le long d'une courbe fermée lisse par morceaux $L \subset G$ qui contourne le compact donné $Q \subset G$ dans le sens positif. Puisque la fonction $\frac{1}{(\zeta - z)^{m+1}}$, pour $\zeta \in L$, $z \in Q$, est majorée en module par la quantité $\frac{1}{d^{m+1}}$, où $d = \inf |\zeta - z|$, et vu que la suite $f_n(\zeta)$ converge uniformément sur le contour L vers $f(\zeta)$, la suite $\frac{f_n(\zeta)}{(\zeta - z)^{m+1}}$ converge vers $\frac{f(\zeta)}{(\zeta - z)^{m+1}}$ uniformément par rapport au paramètre $z \in Q$ sur le contour L . On peut donc appliquer 10.22.b, et le théorème est démontré.

10.37.a. On va établir la réciproque du théorème 10.35:

Théorème. La somme d'une série de puissances

$$f(z) = \sum_0^{\infty} a_n (z - z_0)^n \quad (1)$$

est une fonction analytique à l'intérieur du cercle $G = \{z : |z - z_0| < r\}$ de convergence de la série, et la formule suivante est valable partout à l'intérieur de G :

$$f'(z) = a_1 + 2a_2(z - z_0) + 3a_3(z - z_0)^2 + \dots$$

Démonstration. D'après le théorème 6.64.a, la série (1) converge uniformément dans tout cercle G' inclus dans le cercle de convergence G et ayant le même centre.

D'autre part, n'importe quel compact $Q \subset G$ est contenu dans un tel cercle G' . En effet, si l'on suppose le contraire, on est en mesure de trouver, dans le compact Q , une suite de points $z_1, z_2, z_3, \dots$ avec $|z_n - z_0| \rightarrow r$, et le compact Q contient n'importe quelle valeur d'adhérence z_ω de cette suite. Or, pour z_ω on a $|z_\omega - z_0| = r$, ce qui contredit la condition $Q \subset G$.

On peut donc appliquer le théorème 10.36.b. Les fonctions $(z - z_0)^n$ étant analytiques, la série (1) représente bien, d'après 10.36.b, une fonction analytique à l'intérieur de G .

La deuxième partie du théorème résulte de la deuxième partie de 10.36.b.

b. En particulier, on a

$$(e^z)' = \left(1 + z + \frac{z^2}{2!} + \dots\right)' = 1 + z + \frac{z^2}{2!} + \dots = e^z, \quad (2)$$

$$(\sin z)' = \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots\right)' = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots = \cos z, \quad (3)$$

$$(\cos z)' = \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots\right)' = -z + \frac{z^3}{3!} - \frac{z^5}{5!} + \dots = -\sin z, \quad (4)$$

$$(\operatorname{sh} z)' = \left(z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots\right)' = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \operatorname{ch} z, \quad (5)$$

$$(\operatorname{ch} z)' = \left(1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots\right)' = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \operatorname{sh} z. \quad (6)$$

D'ailleurs, on pourrait obtenir les formules (3)-(6) à partir de la formule (2) en se servant des expressions connues des fonctions trigonométriques et hyperboliques par l'exponentielle (8.63 et 8.72).

c. Le rayon de convergence de la série (1) est lié, d'une façon naturelle, à l'analyticité de la fonction $f(z)$. Notamment, le cercle de convergence de la série (1) est le plus grand cercle centré en z_0 dans lequel la fonction $f(z)$ est analytique, puisque la fonction $f(z)$ est représentée par sa série de Taylor dans tout cercle où elle est analytique (10.35).

Ainsi, le rayon de convergence de la série

$$f(z) = \frac{1}{1+z^2} = 1 - z^2 + z^4 - z^6 + \dots \quad (7)$$

vaut 1, et les singularités de la fonction $f(z)$ les plus proches du point $z_0 = 0$ se trouvent en $z = \pm i$, à la distance 1 du point z_0 . En restant dans le domaine réel, il est difficile d'expliquer, à partir des propriétés de la fonction $f(z)$, la divergence de la série (7) pour $z = \pm 1$.

d. Un point $z = a$ s'appelle *zéro* ou *racine* d'une fonction $w = f(z)$ analytique dans un domaine $G \ni a$ si $f(a) = 0$. Plus précisément, ce point est appelé *zéro (racine) de multiplicité k* si $f(a) = f'(a) = \dots = f^{(k-1)}(a) = 0$, $f^{(k)}(a) \neq 0$. Ainsi, au voisinage de son zéro de multiplicité k une fonction analytique $f(z)$ est développée en série de Taylor qui commence par $(z-a)^k$:

$$f(z) = \frac{f^{(k)}(a)}{k!} (z-a)^k + \frac{f^{(k+1)}(a)}{(k+1)!} (z-a)^{k+1} + \dots$$

On peut aussi bien noter

$$f(z) = (z-a)^k g(z),$$

où

$$g(z) = \frac{f^{(k)}(a)}{k!} + \frac{f^{(k+1)}(a)}{(k+1)!} (z-a) + \dots$$

est analytique (en tant que somme d'une série de puissances convergente) au voisinage du point a , donc aussi (en tant que rapport de deux fonctions analytiques $f(z)$ et $(z-a)^k$) dans tout le domaine G . De plus, on a $g(a) = \frac{1}{k!} f^{(k)}(a) \neq 0$.

10.38. Une fonction analytique dans tout le plan C s'appelle *fonction analytique entière*. D'après 10.35, toute fonction entière $f(z)$ se développe en une série de Taylor

$$f(z) = \sum_0^{\infty} a_n z^n \quad (z_0 = 0) \quad (1)$$

qui converge dans tout le plan.

Théorème (Liouville). *Une fonction analytique entière $f(z)$ qui est bornée en module dans tout le plan C est une constante.*

Démonstration. Supposons que

$$|f(z)| \leq A$$

pour tous les $z \in C$. Considérons l'expression 10.35 (2) des coefficients de la série en choisissant pour contour Λ la circonférence d'un rayon R centrée en l'origine des coordonnées:

$$a_n = \frac{1}{2\pi i} \oint_{\Lambda} \frac{f(\zeta) d\zeta}{\zeta^{n+1}}. \quad (2)$$

Vu 10.21 (5), on a

$$|a_n| \leq \frac{1}{2\pi} \cdot \frac{A}{R^{n+1}} \cdot 2\pi R = \frac{A}{R^n}.$$

La quantité R , dans cette inégalité, étant arbitrairement grande on obtient

$$a_n = 0 \quad \text{pour } n = 1, 2, \dots$$

donc

$$f(z) = a_0 = \text{const},$$

ce qu'il fallait démontrer.

10.39.a. Théorème de l'unicité d'une fonction analytique. Si une fonction $\varphi(z)$ analytique dans un domaine connexe G est nulle sur une suite de points $z_1, z_2, \dots$ qui converge vers un point $z_0 \in G$, alors $\varphi(z) \equiv 0$ dans le domaine G .

Démonstration. En vertu du théorème 10.35, la fonction $\varphi(z)$, dans tout cercle $V = \{z: |z - z_0| < r\} \subset G$ est la somme d'une série de puissances

$$\varphi(z) = c_0 + c_1(z - z_0) + c_2(z - z_0)^2 + \dots \quad (1)$$

Montrons d'abord que $\varphi(z) \equiv 0$ dans le cercle V . Il suffit de prouver que tous les nombres $c_0, c_1, c_2, \dots$ sont nuls. A côté de la série (1), considérons les séries de puissances

$$\varphi_1(z) = c_1 + c_2(z - z_0) + c_3(z - z_0)^2 + \dots,$$

$$\varphi_2(z) = c_2 + c_3(z - z_0) + c_4(z - z_0)^2 + \dots,$$

$$\dots \dots \dots$$

Chaque série suivante s'obtient de la précédente dont on retranche son terme constant et qu'on divise ensuite par $(z - z_0)$. Par conséquent, toutes les séries convergent dans le même cercle que la série (1), et les fonctions $\varphi_m(z)$ sont continues, d'après 6.64.b, pour $z = z_0$. On a ensuite par récurrence:

$$\varphi(z_n) = 0, \quad c_0 = \varphi(z_0) = \lim_{n \rightarrow \infty} \varphi(z_n) = 0;$$

$$\varphi_1(z_n) = \frac{\varphi(z_n) - \varphi(z_0)}{z_n - z_0} = 0, \quad c_1 = \varphi_1(z_0) = \lim_{n \rightarrow \infty} \varphi_1(z_n) = 0;$$

$$\dots \dots \dots$$

$$\varphi_m(z_n) = \frac{\varphi_{m-1}(z_n) - \varphi_{m-1}(z_0)}{z_n - z_0} = 0, \quad c_m = \varphi_m(z_0) = \lim_{n \rightarrow \infty} \varphi_m(z_n) = 0;$$

$$\dots \dots \dots$$

Définitivement, tous les c_l sont nuls.

On a démontré que $\varphi(z) \equiv 0$ dans un certain cercle V centré en z_0 . Soit maintenant w un point quelconque du domaine G . Comme G est connexe, il existe un chemin lisse par morceaux $L \subset G$ qui lie les points z_0 et w .

Posons $d = \inf \rho(z)$, où $\rho(z)$ est le rayon du plus grand cercle possible centré en un point $z \in L$ et inclus dans G ; le nombre d est positif en vertu de la compacité de L . On a vu plus haut que $\varphi(z) \equiv 0$ dans le cercle V_0 de rayon d centré en z_0 . En déplaçant le centre du cercle de rayon d le long de L , on voit que $\varphi(z) \equiv 0$ à l'intérieur du cercle pour toute position de son centre sur la ligne L . En arrivant au point w on obtient $\varphi(w) = 0$. Le théorème est démontré.

b. Théorème. *Si deux fonctions $f(z)$ et $g(z)$ analytiques dans un domaine connexe G se confondent sur une suite de points $z_1, \dots, z_n, \dots$ qui converge vers un point $z_0 \in G$, alors $f(z) \equiv g(z)$ dans le domaine G .*

Démonstration. Posons $f(z) - g(z) = \varphi(z)$; la fonction $\varphi(z)$ est analytique dans G de même que $f(z)$ et $g(z)$ et satisfait à l'hypothèse du théorème a. En appliquant a on voit que $\varphi(z) \equiv 0$ dans le domaine G , donc $f(z) = g(z)$ partout dans G . Le théorème est démontré.

c. La propriété b est à la base de l'idée de prolongement analytique d'une fonction. Soient $f_1(z)$ une fonction analytique dans un domaine G_1 et $f_2(z)$ une fonction analytique dans un domaine G_2 dont l'intersection G_0 avec G_1 est non vide et connexe. Supposons qu'il y ait, dans G_0 , une suite de points $z_1, z_2, \dots, z_n, \dots$ convergeant vers un point $z_0 \in G_0$ et telle que $f_1(z_n) = f_2(z_n)$ ($n = 1, 2, \dots$). Alors, d'après ce qu'on a vu plus haut, les fonctions $f_1(z)$ et $f_2(z)$ se confondent dans tout le domaine G_0 ; la fonction $f(z)$ qui vaut $f_1(z)$ dans G_1 et $f_2(z)$ dans G_2 est définie d'une façon univoque dans le domaine $G = G_1 \cup G_2$ et analytique dans le domaine G . On dit alors que la fonction $f(z)$ est un *prolongement analytique de la fonction $f_1(z)$ de G_1 dans G* . Il résulte de ce qu'on a dit qu'il ne peut exister qu'un seul prolongement analytique de $f_1(z)$ du domaine G_1 dans le domaine G .

d. Exemple. La fonction $f_1(z) = \sum_{n=0}^{\infty} z^n$ est définie (comme somme de la série) et analytique dans le cercle $|z| < 1$. La fonction $f(z) = \frac{1}{1-z}$ qui est définie partout sauf en $z = 1$ représente le prolongement analytique de la fonction $f_1(z)$ dans tout le plan C dépourvu du point $z = 1$.

e. Théorème. *Si toutes les dérivées $f^{(k)}(a)$ d'une fonction $f(z)$ analytique dans un domaine connexe G s'annulent au point a ($k = 0, 1, 2, \dots$), alors $f(z) \equiv 0$ dans G .*

Démonstration. Soit un cercle $V = \{z: |z - a| \leq r\}$ contenu entièrement dans le domaine G . D'après le théorème 10.35.c, la fonction $f(z)$ est nulle dans tout le cercle, et l'on applique le théorème 10.39.a.

f. Exemple: fonctions analytiques réelles. Une fonction $f(z)$ analytique dans un domaine connexe G qui contient des points de l'axe réel s'appelle *fonction analytique réelle* si toutes ses valeurs sur l'axe réel sont réelles.

Comme les valeurs que la dérivée d'une fonction analytique $f(z)$ prend sur l'axe réel peuvent être calculées d'après les valeurs prises par la fonction $f(z)$ sur l'axe réel, on voit que la dérivée $f'(z)$ d'une fonction analytique réelle est une fonction analytique réelle. De même, toutes les autres dérivées $f''(z)$, $f'''(z)$, ... sont aussi des fonctions analytiques réelles. Le développement de Taylor de la fonction $f(z)$ au voisinage d'un point réel z_0

$$f(z) = \sum_0^{\infty} a_n (z - z_0)^n$$

a les coefficients réels $a_n = \frac{1}{n!} f^{(n)}(z_0)$. Nous avons donc, dans le cercle de convergence de cette série:

$$\overline{f(\bar{z})} = \overline{\sum_0^{\infty} a_n (\bar{z} - z_0)^n} = \sum_0^{\infty} a_n (z - z_0)^n = f(z).$$

Théorème. Une fonction analytique réelle $f(z)$ satisfait à la condition

$$\overline{f(\bar{z})} = f(z)$$

pour tout point $z \in G$ tel que $\bar{z} \in G$.

Démonstration. Les points qui appartiennent à G avec leurs complexes conjugués forment un domaine symétrique par rapport à l'axe réel. Il suffit donc de discuter le cas où G est symétrique par rapport à l'axe réel. Désignons par G^+ la partie de G contenue dans le demi-plan supérieur, par G^- la partie contenue dans le demi-plan inférieur. Définissons, dans le domaine G^+ , une nouvelle fonction

$$f_1(z) = \overline{f(\bar{z})}$$

en nous servant des valeurs de $f(z)$ pour $z \in G^-$. Pour tout $z_0 \in G^+$, le développement de Taylor

$$f(z) = \sum_0^{\infty} a_n (z - z_0)^n$$

fournit le développement de la fonction $f_1(z)$ en la série de puissances

$$f_1(z) = \overline{f(\bar{z})} = \overline{\sum_0^{\infty} a_n (\bar{z} - \bar{z}_0)^n} = \sum_0^{\infty} \bar{a}_n (z - z_0)^n,$$

ce qui veut dire que la fonction $f_1(z)$ est analytique dans G^+ . Or, au voisinage de l'axe réel, la fonction $f_1(z)$ se confond, on l'a vu, avec $f(z)$. En appliquant le théorème de prolongement analytique, on voit que $f_1(z) \equiv f(z)$ partout dans G^+ . En remplaçant G^+ par G^- et vice versa on obtient que l'identité $\overline{f(\bar{z})} \equiv f(z)$ a également lieu partout dans G^- . Sur l'axe réel, on a évidemment $\overline{f(\bar{z})} = f(z)$. Donc, $\overline{f(\bar{z})} \equiv f(z)$ partout dans G , ce qu'il fallait démontrer.

§ 10.4. Résidus et points singuliers isolés

10.41. Résidus. Si une fonction $f(z)$ est analytique dans un domaine connexe G contenant un contour L fermé lisse par morceaux, alors, d'après le théorème de Cauchy 10.32, on a

$$\frac{1}{2\pi i} \oint_L f(\zeta) d\zeta = 0.$$

On a souvent affaire à une intégrale de la forme

$$\frac{1}{2\pi i} \oint_L f(\zeta) d\zeta, \quad (1)$$

la fonction $f(z)$ ayant des points singuliers à l'intérieur du domaine borné de frontière L . Une telle intégrale n'est en général pas nulle. L'intégrale de Cauchy 10.33 (5) en est un exemple :

$$\frac{1}{2\pi i} \oint_L \frac{\varphi(\zeta) d\zeta}{\zeta - z} = \varphi(z), \quad (2)$$

où $\varphi(z)$ est analytique dans le domaine simplement connexe G qui contient le chemin L . Considérons le cas où la fonction $f(z)$ dans (1) possède, dans le domaine G contenant le contour L , un seul point singulier $z = a$ (singulier en ce sens que ou bien la fonction $f(z)$ n'y est pas définie du tout, ou bien elle n'y est pas dérivable par rapport au domaine G). La valeur de l'intégrale (1) ne dépend pas, dans ce cas, du choix du contour L (à condition qu'il contourne le point a une fois dans le sens positif). Cette valeur est appelée *résidu* de la fonction $f(z)$ au point $z = a$ et désignée comme suit : $\text{Res}[f(z)]_{z=a}$.

10.42. Voici quelques exemples.

a. L'intégrale

$$I = \frac{1}{2\pi i} \oint_L \frac{g(\zeta)}{P(\zeta)} d\zeta,$$

où $P(z)$ est un polynôme dont $z = a$ est une racine simple ($P(a) = 0$, $P'(a) \neq 0$) et qui n'a pas d'autres racines dans le domaine G , et $g(z)$ est une fonction analytique dans G . On a

$$P(z) = b_1(z-a) + b_2(z-a)^2 + \dots = (z-a) P_1(z),$$

où $P_1(z) = b_1 + b_2(z-a) + \dots$ est un polynôme qui vaut $b_1 = P'(a) \neq 0$ pour $z = a$ et qui n'a pas de racines pour $z \in G$. D'après la formule de Cauchy 10.41 (2), on a

$$I = \frac{1}{2\pi i} \oint_L \frac{g(\zeta)}{P(\zeta)} d\zeta = \frac{1}{2\pi i} \oint_L \frac{g(\zeta) d\zeta}{(\zeta-a) P_1(\zeta)} = \frac{\kappa(a)}{P_1(a)} = \frac{g(a)}{P'(a)},$$

et le résidu de la fonction $f(z) = \frac{g(z)}{P(z)}$ pour $z = a$ est trouvé :

$$\text{Res}[f(z)]_{z=a} = I = \frac{g(a)}{P'(a)}. \quad (1)$$

b. L'intégrale

$$I = \frac{1}{2\pi i} \oint_L \frac{g(\zeta)}{P(\zeta)} d\zeta,$$

où le polynôme $P(z)$, contrairement à celui de l'exemple précédent, possède en $z = a$ une racine de multiplicité m , de sorte que

$$P(z) = b_m(z-a)^m + b_{m+1}(z-a)^{m+1} + \dots, \quad b_m \neq 0.$$

On a

$$P(z) = (z-a)^m P_m(z),$$

où

$$P_m(z) = b_m + b_{m+1}(z-a) + \dots, \quad P_m(a) = b_m = \frac{P^{(m)}(a)}{m!} \neq 0.$$

Développons la fonction $\frac{g(z)}{P_m(z)}$ en série de Taylor suivant les puissances de $(z-a)$:

$$\frac{g(z)}{P_m(z)} = c_0 + c_1(z-a) + c_2(z-a)^2 + \dots$$

Alors, on a

$$\begin{aligned} \frac{g(z)}{P(z)} &= \frac{g(z)}{(z-a)^m P_m(z)} = \frac{c_0}{(z-a)^m} + \frac{c_1}{(z-a)^{m-1}} + \dots \\ &\quad \dots + \frac{c_{m-1}}{z-a} + c_m + c_{m+1}(z-a) + \dots \end{aligned}$$

Après l'intégration le long de L , il ne restera que le terme contenant le coefficient c_{m-1} (10.33.a). D'après 10.35.b, il vaudra

$$I = \frac{1}{2\pi i} \oint_L \frac{g(\zeta) d\zeta}{P(\zeta)} = c_{m-1} = \frac{1}{(m-1)!} \left(\frac{g(z)}{P_m(z)} \right)_{z=a}^{(m-1)} = \\ = \frac{1}{(m-1)!} \left(\frac{g(z)(z-a)^m}{P(z)} \right)_{z=a}^{(m-1)}.$$

Définitivement,

$$\text{Res} \left[\frac{g(z)}{P(z)} \right]_{z=a} = \frac{1}{(m-1)!} \left[\frac{g(z)(z-a)^m}{P(z)} \right]_{z=a}^{(m-1)} \quad (2)$$

10.43. Lorsqu'il faut calculer l'intégrale 10.41 (1) le long d'une courbe fermée contournant non pas un seul, mais plusieurs points

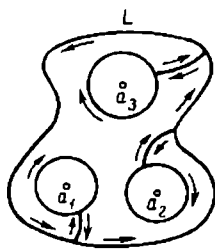


Fig. 10.6.

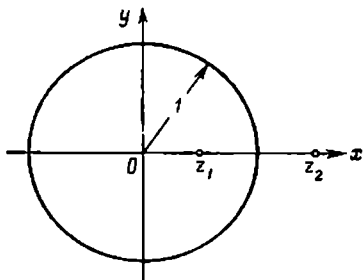


Fig. 10.7.

singuliers $a_1, \dots, a_m$ de la fonction $f(z)$, on peut se servir de la relation

$$\frac{1}{2\pi i} \oint_L f(\zeta) d\zeta = \frac{1}{2\pi i} \oint_{L_1} f(\zeta) d\zeta + \dots + \frac{1}{2\pi i} \oint_{L_m} f(\zeta) d\zeta = \\ = \sum_{j=1}^m \text{Res} [f(z)]_{z=a_j}, \quad (1)$$

où $L_1, \dots, L_m$ sont les courbes fermées lisses par morceaux chacune desquelles contournent un seul point singulier. La démonstration de cette égalité (appelée *théorème fondamental des résidus*) est évidente si l'on tient compte du théorème de Cauchy et de la fig. 10.6 qui montre le chemin initial L , les chemins $L_1, \dots, L_m$ et les chemins auxiliaires parcourus chacun deux fois dans les sens opposés.

Exemple. Calculons l'intégrale

$$\oint_{|z|=1} \frac{dz}{z^2 - 2pz + 1}, \quad p > 1.$$

Le produit de deux racines du dénominateur vaut 1. L'une d'elles, z_1 , se trouve à l'intérieur, l'autre, z_2 , à l'extérieur du cercle unité (fig. 10.7). Dans la formule des racines du trinôme du second degré

$$z_{1,2} = p \pm \sqrt{p^2 - 1}$$

pour $p > 1$, le signe moins correspond à la racine z_1 , le signe plus à z_2 . En z_1 , le dénominateur a une racine simple, et l'on a, d'après la formule 10.42 (1) :

$$\oint_{|z|=1} \frac{dz}{z^2 - 2pz + 1} = 2\pi i \frac{1}{2(z-p)} \Big|_{z=z_1} = \frac{-i\pi}{\sqrt{p^2 - 1}}.$$

10.44. Résidu logarithmique. Soit une intégrale de la forme

$$N = \frac{1}{2\pi i} \oint_L \frac{f'(z)}{f(z)} dz, \quad (1)$$

où la fonction $f(z)$ est analytique dans un domaine G contenant le contour fermé L et ne s'annule pas sur ce contour. Si elle ne s'annule pas non plus à l'intérieur du contour L , alors l'intégrale (1) est nulle d'après le théorème de Cauchy 10.32. Si la fonction $f(z)$ a des zéros à l'intérieur de L , alors, d'après le théorème d'unicité 10.39.a, ils sont en nombre fini; désignons-les par $z_1, z_2, \dots, z_m$, et leur multiplicités respectives par $k_1, k_2, \dots, k_m$. Conformément à 10.43, l'intégrale N est la somme des intégrales analogues

$$N = \sum_{j=1}^m N_j, \quad N_j = \frac{1}{2\pi i} \oint_{L_j} \frac{f'(z)}{f(z)} dz,$$

où chaque contour L_j ne contient à son intérieur qu'un seul zéro z_j de multiplicité k_j . Pour abréger la notation, posons $z_j = a$, $k_j = k$. Au voisinage du point a , on a

$$\begin{aligned} f(z) &= c_k (z-a)^k + c_{k+1} (z-a)^{k+1} + \dots, \quad c_k \neq 0, \\ f'(z) &= k c_k (z-a)^{k-1} + (k+1) c_{k+1} (z-a)^k + \dots, \\ \frac{f'(z)}{f(z)} &= \frac{1}{z-a} \cdot \frac{k c_k + (k+1) c_{k+1} (z-a) + \dots}{c_k + c_{k+1} (z-a) + \dots} = \frac{\varphi(z)}{z-a}, \end{aligned}$$

où $\varphi(z)$ est une fonction analytique au voisinage du point a qui prend en a la valeur $\varphi(a) = k$. Donc,

$$\frac{1}{2\pi i} \oint_{L_j} \frac{f'(z)}{f(z)} dz = \frac{1}{2\pi i} \oint_{L_j} \frac{\varphi(z)}{z-a} dz = \varphi(a) = k.$$

Définitivement

$$N = \sum_{j=1}^m N_j = \sum_{j=1}^m k_j,$$

c'est-à-dire que l'intégrale N est un entier positif égal à la somme des multiplicités de toutes les racines de la fonction $f(z)$ qui sont à l'intérieur du contour L .

L'intégrale N s'appelle *résidu logarithmique de la fonction $f(z)$ par rapport au contour L* . L'origine de ce terme sera éclaircie par la suite (10.57).

10.45. Le présent numéro ainsi que plusieurs numéros suivants sont consacrés à l'« analyse qualitative » du comportement d'une fonction au voisinage d'un point singulier isolé. C'est le développement de la fonction considérée en la série, dite de *Laurent*, suivant les puissances positives et négatives de $z - a$ qui en sert d'instrument.

Soit $w = f(z)$ une fonction analytique dans une couronne circulaire $Q = \{r_0 < |z - a| < R_0\}$, où $0 \leq r_0 < R_0 \leq \infty$ (fig. 10.8). Choisissons les nombres r et R de façon à avoir $r_0 < r < R < R_0$. Alors, si l'on applique le théorème de Cauchy au domaine simplement connexe que l'on obtient en coupant la couronne suivant un rayon ne passant pas par z , on aura, pour $r < |z - a| < R$:

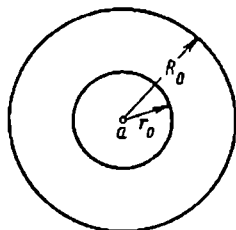


Fig. 10.8.

$$f(z) = \frac{1}{2\pi i} \oint_{|z-a|=R} \frac{f(\zeta) d\zeta}{\zeta - z} - \frac{1}{2\pi i} \oint_{|z-a|=r} \frac{f(\zeta) d\zeta}{\zeta - z}.$$

Pour la première intégrale, on utilise le développement

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - a) - (z - a)} = \frac{1}{(\zeta - a) \left[1 - \frac{z - a}{\zeta - a} \right]} = \frac{1}{\zeta - a} \sum_{n=0}^{\infty} \left(\frac{z - a}{\zeta - a} \right)^n,$$

pour la deuxième, le développement

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - a) - (z - a)} = -\frac{1}{z - a} \cdot \frac{1}{1 - \frac{\zeta - a}{z - a}} = -\sum_{n=0}^{\infty} \frac{(\zeta - a)^n}{(z - a)^{n+1}}.$$

Selon le critère de Weierstrass (cf. 10.35.a), les deux développements convergent uniformément sur les contours correspondants d'intégration. Donc, en appliquant 10.22.a. on a

$$f(z) = \sum_0^{\infty} c_n (z - a)^n + \sum_{-1}^{-\infty} c_n (z - a)^n, \quad (1)$$

où

$$c_n = \begin{cases} \frac{1}{2\pi i} \oint_{|z-a|=R} \frac{f(\zeta) d\zeta}{(\zeta-a)^{n+1}} & \text{pour } n \geq 0, \\ \frac{1}{2\pi i} \oint_{|z-a|=r} \frac{f(\zeta) d\zeta}{(\zeta-a)^{n+1}} & \text{pour } n < 0. \end{cases}$$

D'ailleurs, la fonction à intégrer $\frac{f(\zeta)}{(\zeta-a)^{n+1}}$ étant analytique dans la couronne Q , on peut intégrer le long de n'importe quelle circonférence $|z-a| = \rho$, $\rho \in (r_0, R_0)$ et même le long de toute courbe fermée lisse par morceaux qui contourne le point a une fois dans le sens positif sans sortir de la couronne Q .

En particulier, si $r_0 = 0$, alors le coefficient c_{-1} est bien le résidu de la fonction $f(z)$ au point $z = a$ (10.41).

La notation abrégée de la somme (1) est

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n. \quad (2)$$

On peut aussi bien considérer cette expression comme limite de la somme $\sum_{n=p}^q c_n (z-a)^n$ lorsque $p \rightarrow -\infty$, $q \rightarrow \infty$ indépendamment l'un de l'autre, c.-à-d. comme somme d'une série bilatérale (6.48.a); ceci résulte de la convergence des deux séries dans le second membre de (1) et du théorème 6.48.b.

La série (1) ou (2) s'appelle *série de Laurent* pour la fonction $w = f(z)$ dans la couronne Q . D'après le théorème 10.22.b, elle converge uniformément pour $r \leq |z-a| \leq R$. Elle est définie d'une façon univoque pour toute fonction $f(z)$, parce que, quels que soient $m = 0, \pm 1, \pm 2, \dots$, $r_0 < \rho < R_0$, il résulte de (1) que

$$\int_{|z-a|=\rho} f(z) (z-a)^{-m} dz = \sum_{n=-\infty}^{\infty} c_n \int_{|z-a|=\rho} (z-a)^{-m+n} dz = 2\pi i c_{m-1}.$$

Le terme $f^+(z) = \sum_0^{\infty} c_n (z-a)^n$ est appelé *partie régulière* de la série de Laurent, le terme $f^-(z) = \sum_{-1}^{-\infty} c_n (z-a)^n$ *partie principale*. Les nombres c_n s'appellent *coefficients de Laurent*.

La partie régulière est une série suivant les puissances positives de $z-a$ qui converge pour $|z-a| = R$, donc aussi dans tout le cercle $|z-a| \leq R_0$ (6.62), y compris la partie intérieure $|z-a| \leq r_0$ qui n'est pas contenue dans la couronne Q .

La partie principale est une série suivant les puissances négatives de $z - a$ qui converge pour $|z - a| = r$. Si l'on pose $z - a = \frac{1}{\xi}$, on obtient une série suivant les puissances positives de ξ qui converge pour $|\xi| = \frac{1}{r}$, donc aussi (6.62) pour $|\xi| \leq \frac{1}{r}$; par conséquent, la partie principale de la série de Laurent converge pour tous les $|z - a| > r_0$, y compris la partie extérieure $|z - a| \geq R_0$ qui n'est pas contenue dans la couronne Q .

10.46. Points singuliers isolés. Soit $w = f(z)$ une fonction analytique dans tout le cercle $|z - a| < R_0$ à l'exception du point au centre a . Dans le domaine $0 < |z - a| < R_0$, la fonction $f(z)$ se développe en série de Laurent

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n. \quad (1)$$

La structure de cette série détermine l'appellation du point singulier isolé a .

a. Si tous les coefficients de Laurent c_n avec $n < 0$ sont nuls, le point a s'appelle *point singulier éliminable* de la fonction $f(z)$.

La fonction $f(z)$ se développe alors, dans le domaine $0 < |z - a| < R_0$, en la série

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n$$

dont la somme est une fonction analytique dans tout le cercle $|z - a| < R_0$, y compris le point $z = a$, et prend la valeur c_0 pour $z = a$. Donc, si l'on complète la définition de la fonction $f(z)$ en posant $f(a) = c_0$, on aboutit à une fonction analytique dans le cercle $|z - a| < R_0$ sans aucune singularité au point $z = a$. Ceci explique le terme « point singulier éliminable ».

b. Si tous les coefficients de Laurent c_n avec $n < -m$ et $m > 0$ sont nuls et que $c_{-m} \neq 0$, alors le point a est appelé *pôle d'ordre m de la fonction $f(z)$* .

c. S'il y a des coefficients de Laurent c_n non nuls avec n négatifs et arbitrairement grands en module, alors le point a est appelé *point singulier essentiel* de la fonction $f(z)$.

d. Pour la raison de généralité, complétons les plans complexes des variables z et w par les points impropres $z = \infty$, $w = \infty$ respectivement (« points à l'infini »). Nous désignons par le symbole $z \rightarrow \infty$ la direction dans le plan des z formée des ensembles $A_r = \{z: |z| > r\}$, où r parcourt tous les nombres positifs (ou, au moins, tous les nombres supérieurs à un r_0); le symbole $w \rightarrow \infty$ a un sens analogue (4.73.f). Les points z et w distincts des $z = \infty$ et $w = \infty$ que l'on vient de définir sont appelés *points finis*. Soit

une direction S dans le plan des z ; nous dirons que pour une fonction donnée $w = f(z)$ la relation $\lim_{z \rightarrow \infty} f(z) = \infty$ a lieu si, quel que soit un nombre r , il existe un ensemble $A_r \in S$ en tout point duquel on a

$$|f(z)| > r.$$

De la sorte, on attribue un sens exact aux expressions

$$\lim_{z \rightarrow a} f(z) = \infty, \quad \lim_{z \rightarrow \infty} f(z) = b, \quad \lim_{z \rightarrow \infty} f(z) = \infty,$$

a et b étant des points à distance finie. Par exemple, on a

$$\lim_{z \rightarrow a} \frac{1}{(z-a)^m} = \infty;$$

$\lim_{z \rightarrow a} \varphi(z) f(z) = \infty$ lorsque $\lim_{z \rightarrow a} \varphi(z) = b \neq 0$; $\lim_{z \rightarrow a} f(z) = \infty$;

$$\lim_{z \rightarrow a} \varphi(z) = \infty \text{ si, et seulement si, } \lim_{z \rightarrow a} \frac{1}{\varphi(z)} = 0.$$

D'une façon analogue à 4.45, on dit que le point $w = \infty$ est une *valeur d'adhérence de la fonction $f(z)$ suivant une direction S* lorsque, pour tout r , il existe dans chacun des ensembles $A_r \in S$ un point z , tel que

$$|f(z_r)| > r.$$

Si une fonction $f(z)$ n'est pas bornée suivant la direction S , elle a le point $w = \infty$ pour sa valeur d'adhérence.

e. Considérons le cas où une fonction $f(z)$ est définie et analytique dans un domaine $G = \{z : |z| > R\}$. En vertu du théorème général 10.45, elle se développe, dans le domaine G , en sa série de Laurent

$$f(z) = \sum_{n=-\infty}^{\infty} c_n z^n. \quad (2)$$

L'application $z = \frac{1}{\zeta}$ transforme le domaine G en le domaine $H = \{\zeta : 0 < |\zeta| < \frac{1}{R}\}$ et la fonction $f(z)$ en la fonction

$$\varphi(\zeta) = \sum_{n=-\infty}^{\infty} c_n \zeta^{-n}. \quad (3)$$

On dit que la fonction $f(z)$ a, au point ∞ , un *point singulier éliminable*, un *pôle* ou un *point singulier essentiel* si la fonction $\varphi(\zeta)$, pour $\zeta = 0$, a un point singulier du même type. Autrement dit, le point $z = \infty$ est un *point singulier éliminable* pour la fonction $f(z)$ si le développement (2) ne contient aucune puissance d'exposant positif; il est un *pôle d'ordre m* si les coefficients $a_{m+1}, a_{m+2}, \dots$,

dans le développement (2), sont nuls, tandis que $a_m \neq 0$; il est un *point singulier essentiel* si le développement (2) comprend des puissances d'exposants arbitrairement grands.

A présent, on va étudier le comportement de la fonction $f(z)$ à la proximité d'un point singulier dans les trois cas signalés.

10.47. Si, dans le développement 10.46 (1), on a $c_n = 0$ pour $n < -m$, $m > 0$, de sorte que

$$f(z) = \sum_{n=-m}^{\infty} c_n (z-a)^n,$$

alors on pose

$$\varphi(z) = (z-a)^m f(z) = c_{-m} + c_{-m+1}(z-a) + \dots$$

La fonction $\varphi(z)$ a un point singulier éliminable en $z=a$ et, dans la condition supplémentaire $\varphi(a) = c_{-m}$, devient analytique dans un cercle $|z-a| < R_0$ et non nulle en $z=a$. Il en résulte que

$$f(z) = \frac{\varphi(z)}{(z-a)^m}.$$

Étant donné $\varphi(a) \neq 0$, la fonction $f(z)$ tend vers l'infini pour $z \rightarrow a$ (10.46.d).

Les fonctions à intégrer des exemples 10.42.a, b, dont on a calculé les résidus, sont de ce type.

10.48. Supposons maintenant que a soit un point singulier essentiel de la fonction $f(z)$.

Montrons tout d'abord que la fonction $f(z)$ n'est pas bornée au voisinage du point $z=a$. Si $f(z)$ était bornée au voisinage de $z=a$, il en serait de même de sa partie principale $f^-(z)$, puisque la partie régulière $f^+(z)$ y est bien bornée. La série

$$f^-(z) = \sum_{n=1}^{\infty} c_n (z-a)^n$$

converge pour tout $z \neq a$ (10.45). Donc, en remplaçant $z-a$ par $\frac{1}{\zeta}$, on obtient une série

$$\varphi(\zeta) = \sum_{n=1}^{\infty} c_n \zeta^n$$

qui converge dans tout le plan des ζ . Si la fonction $f^-(z)$ était bornée pour $0 < |z-a| \leq \varepsilon$, la fonction $\varphi(\zeta)$ le serait pour $|\zeta| \geq \frac{1}{\varepsilon}$.

Or, la fonction $\varphi(\zeta)$ est bornée pour $|\zeta| < \frac{1}{\varepsilon}$ (puisque'elle est continue), donc $\varphi(\zeta)$ serait bornée dans tout le plan des ζ . D'après le théorème de Liouville (10.38), $\varphi(\zeta)$ serait constante, donc, étant donné $\varphi(0) = 0$, on aurait $\varphi(\zeta) = 0$, d'où $c_n = 0$ pour tout $n < 0$,

ce qui contredit l'hypothèse. Définitivement, $f(z)$ n'est pas bornée pour $z \rightarrow a$, c'est-à-dire que le point ∞ est une valeur d'adhérence de la fonction $f(z)$ pour $z \rightarrow a$. Soit maintenant A un nombre complexe quelconque. Si la fonction $f(z)$ prend la valeur A dans tout voisinage du point a , alors A est une valeur d'adhérence de $f(z)$ pour $z \rightarrow a$. Si, au contraire, il existe un voisinage du point a dans lequel la fonction $f(z)$ ne prend pas la valeur A , alors la fonction $\varphi(z) = \frac{1}{f(z) - A}$ n'a pas, dans ce voisinage, de points singuliers autres que a lui-même. Si le point a est un pôle ou un point singulier essentiel de la fonction $\varphi(z)$, alors, d'après ce qu'on a vu plus haut, il existe une suite $z_n \rightarrow a$ pour laquelle $\varphi(z_n) \rightarrow \infty$, donc aussi $f(z_n) \rightarrow A$, de sorte que A est une valeur d'adhérence de la fonction $f(z)$ pour $z \rightarrow a$. Envisageons enfin le cas où le point $z = a$ est un point singulier éliminable de la fonction $\varphi(z)$. Il existe alors, comme on l'a vu plus haut, la limite finie

$$b = \lim_{z \rightarrow a} \varphi(z).$$

On sait qu'il existe une suite $z_n \rightarrow a$ pour laquelle $f(z_n) \rightarrow \infty$; par conséquent, on a

$$b = \lim_{z \rightarrow a} \varphi(z) = \lim_{z \rightarrow a} \frac{1}{f(z) - A} = 0.$$

On peut poser

$$\varphi(z) = (z - a)^m \psi(z),$$

où $\psi(a) \neq 0$, $m \geq 1$. La fonction $\frac{1}{\psi(z)}$ est analytique au voisinage du point a et admet le développement de Taylor

$$\frac{1}{\psi(z)} = b_0 + b_1(z - a) + \dots, \quad \text{où} \quad b_0 = \frac{1}{\psi(a)}.$$

Par conséquent,

$$\begin{aligned} f(z) = A + \frac{1}{\varphi(z)} &= A + \frac{1}{(z - a)^m \psi(z)} = \\ &= A + (z - a)^{-m} [b_0 + b_1(z - a) + \dots], \end{aligned}$$

et l'on voit que la partie principale de la série de Laurent pour la fonction $f(z)$ n'a qu'un nombre fini de termes, ce qui contredit l'hypothèse.

Définitivement, pour un point singulier essentiel a , tout nombre A du plan achevé complexe des z est une valeur d'adhérence de la fonction $f(z)$ lorsque $z \rightarrow a^*$.

* Il existe un théorème plus exact de Picard qui dit que, dans tout voisinage de son point singulier essentiel, la fonction $f(z)$ prend réellement un nombre infini de fois n'importe quelle valeur A (à l'exception éventuelle d'une seule). Cf. [5].

10.49. D'autre part, indépendamment de ce qui précède, le comportement de la fonction $f(z)$ au voisinage d'un point singulier isolé admet la classification suivante :

- (1) la fonction $f(z)$ est bornée dans un voisinage du point a ;
- (2) la fonction $f(z)$ tend vers la limite ∞ lorsque $z \rightarrow a$;
- (3) la fonction $f(z)$ a, pour $z \rightarrow a$, au moins deux valeurs d'adhérence dont l'une est ∞ .

Les résultats de 10.46-10.48 sont :

un point singulier éliminable est un point du type (1) ;

un pôle est un point du type (2) ;

un point singulier essentiel est un point du type (3).

La classification 10.46, de même que la classification (1)-(3) ci-dessus, étant complète, on peut formuler les propositions réciproques :

un point du type (1) est un point singulier éliminable ;

un point du type (2) est un pôle ;

un point du type (3) est un point singulier essentiel.

Soulignons que ces propositions ne sont pas triviales. Par exemple, la première dit qu'une fonction analytique et bornée au voisinage d'un point a a une limite finie pour $z \rightarrow a$; elle est basée sur les résultats de l'analyse du comportement d'une fonction au voisinage d'un pôle et d'un point singulier essentiel qu'on a faite dans 10.47-10.48.

Vu 10.46.e, toutes les conclusions obtenues s'étendent au cas d'un point singulier isolé à l'infini. En particulier, si la fonction $f(z)$ est analytique pour tout z tel que $|z| > a$, on est amené aux propositions suivantes : lorsque, dans l'hypothèse formulée, la fonction $w = f(z)$ a une seule valeur d'adhérence $w = \infty$ pour $z \rightarrow \infty$, alors $f(z)$ est un polynôme ; toute fonction entière $f(z)$ qui n'est pas un polynôme a n'importe quel nombre $w \in G$ pour sa valeur d'adhérence lorsque $z \rightarrow \infty$.

§ 10.5. Applications du plan complexe dans lui-même. Fonctions élémentaires

10.51. Application conforme. Soit une fonction du premier degré

$$w = b(z - z_0) + w_0 \equiv f(z), \quad (1)$$

où z_0 , w_0 , $b \neq 0$ sont des nombres complexes. Il est évident que $f(z_0) = w_0$. Ensuite, on a $w - w_0 = b(z - z_0)$, d'où

$$|w - w_0| = |b| \cdot |z - z_0|, \quad (2)$$

$$\arg(w - w_0) = \arg b + \arg(z - z_0) \quad (3)$$

(à un multiple de 2π près). Les relations (2), (3) mettent en relief la nature géométrique de l'application $w = f(z)$: tout cercle $|z - z_0| \leq \rho$ subit l'homothétie de rapport $|b|$ et de centre z_0 , puis la rotation autour de z_0 d'angle $\arg b$ et enfin la translation qui transforme son centre en w_0 .

Soit, à présent, $w = f(z)$ une fonction analytique dans un voisinage $|z - z_0| \leq \rho$ d'un point z_0 et telle que $f'(z_0) = b \neq 0$, $f(z_0) = w_0$. D'après la définition de la dérivée, on a

$$w = w_0 + b(z - z_0) + \varepsilon(z, z_0)(z - z_0), \quad (4)$$

où $\varepsilon(z, z_0) \rightarrow 0$ pour $z \rightarrow z_0$.

L'égalité (4) montre que l'application $w = f(z)$ est représentée, à un infiniment petit du second ordre près, par la formule (1), de sorte qu'elle se réduit approximativement à la rotation d'angle $\arg b$ autour du centre z_0 suivie de l'homothétie de rapport $|b|$ et de la translation plaçant son centre en w_0 . Une application de ce genre est dite *conforme au point* z_0 . Ainsi, une fonction analytique $w = f(z)$ réalise une application conforme en tout point z_0 tel que $f'(z_0) \neq 0$. En particulier, une application conforme transforme deux lignes qui forment un angle α au point z_0 en deux lignes formant le même angle α au point w_0 .

10.52. Théorème. Si $f'(z_0) \neq 0$, il existe un voisinage du point z_0 qui s'applique d'une façon biunivoque sur un voisinage du point w_0 .

L'affirmation est immédiate pour la fonction 10.51 (1), où l'on peut résoudre directement l'équation par rapport à z , mais le cas général demande une démonstration. Soit G un voisinage du point z_0 dans lequel la fonction $f(z) - w_0$ ne s'annule qu'une seule fois, en $z = z_0$; les zéros de la fonction analytique $f(z) - w_0$ étant isolés (10.39.a), un tel voisinage existe bien. Soit L le contour servant de frontière au domaine G . Alors, en vertu de 10.44, on a

$$\frac{1}{2\pi i} \oint_L \frac{f'(\zeta) d\zeta}{f(\zeta) - w_0} = 1,$$

le seul zéro de la fonction $f(z) - w_0$ étant de multiplicité 1.

La fonction de w

$$N(w) = \frac{1}{2\pi i} \oint_L \frac{f'(\zeta) d\zeta}{f(\zeta) - w}$$

est définie pour $w = w_0$ (elle y a 1 pour valeur), aussi bien que pour d'autres valeurs de w proches de w_0 , par exemple, pour $|w - w_0| < \varepsilon$, car le dénominateur est non nul non seulement pour $w = w_0$, mais aussi pour des valeurs assez proches de w . Cette fonction est continue d'après le théorème 10.24. Comme ses valeurs sont entières (10.44), elle vaut 1 pour tous les w proches de w_0 . Ceci veut dire, en

vertu de 10.44, que la fonction $f(z) - w$ s'annule une seule fois à l'intérieur du contour L , c'est-à-dire que, quel que soit w_1 , $|w_1 - w_0| < \varepsilon$, il existe dans G un seul point z_1 tel que $f(z_1) = w_1$. Le théorème est démontré.

10.53. Fonction inverse. Dans l'hypothèse 10.52, l'application $w = f(z)$ étant bijective, il existe une fonction inverse $z = \varphi(w)$ qui réalise l'application d'un voisinage du point w_0 sur un voisinage de z_0 . Montrons que la fonction $z = \varphi(w)$ est analytique. Pour le faire, considérons le rapport

$$\frac{\varphi(w) - \varphi(w_0)}{w - w_0} = \frac{\Delta z}{\Delta w}. \quad (1)$$

Si $w \rightarrow w_0$, on a $z \rightarrow z_0$; si la fonction $\varphi(w)$ avait une autre valeur d'adhérence z_1 pour $w \rightarrow w_0$, on aurait $f(z_0) = w_0 = f(z_1)$, ce qui est impossible comme on l'a vu plus haut. Il existe donc la limite du rapport (1) qui vaut

$$\lim_{w \rightarrow w_0} \frac{\Delta z}{\Delta w} = \frac{1}{\lim_{z \rightarrow z_0} \frac{\Delta w}{\Delta z}} = \frac{1}{f'(z_0)}, \quad (2)$$

ce qui démontre la dérivabilité de la fonction $\varphi(w)$ pour $w = w_0$.

10.54. Fonctions injectives. Une fonction $w = f(z)$ définie et analytique dans un domaine G du plan des z est dite *injective* dans G , si elle applique ce domaine d'une façon biunivoque sur un domaine H du plan des w . On verra (10.59.b) que $f'(z) \neq 0$ pour toute fonction injective. D'après 10.52, toute fonction $w = f(z)$ analytique est injective dans un voisinage du point z_0 , si $f'(z_0) \neq 0$.

10.55. Fonction homographique. Une fonction du premier degré

$$w = b(z - z_0) + w_0 \quad (1)$$

considérée dans 10.51 est, évidemment, injective dans tout le plan des z .

La fonction $w = \frac{1}{z}$ réalise une application bijective et conforme, en tout point à distance finie, du plan achevé des z sur le plan achevé des w . Les points $z = 0$ et $z = \infty$ ont pour images $w = \infty$ et $w = 0$ respectivement. On peut dire que la fonction (1) applique elle aussi le plan achevé des z sur le plan achevé des w en faisant correspondre au point $z = \infty$ le point $w = \infty$.

Une fonction homographique de la forme générale

$$w = \frac{az + b}{cz + d}, \quad \text{où } bc \neq ad, \quad c \neq 0,$$

peut être considérée comme composée d'une fonction du premier degré, de la fonction $\frac{1}{z}$ et encore d'une fonction du premier degré; cela résulte de la représentation

$$w = \frac{a}{c} + \frac{bc - ad}{c(cz + d)}.$$

Par conséquent, la fonction homographique appliquée, d'une façon biunivoque, le plan achevé des z sur le plan achevé des w en faisant correspondre aux points $z = -\frac{d}{c}$ et $z = \infty$ les points $w = \infty$ et $w = \frac{a}{c}$ respectivement.

10.56. Fonction puissance $w = z^n$, $n > 1$, et son inverse. Ici on a, au moins pour $|\arg z| < \frac{\pi}{n}$.

$$|w| = |z|^n; \arg w = n \arg z. \quad (1)$$

On voit que la fonction $w = z^n$ réalise une application bijective de l'intérieur de l'angle $-\alpha < \arg z < \alpha \leq \frac{\pi}{n}$ sur l'intérieur de l'angle

$$-n\alpha < \arg w < n\alpha$$

du plan des w . Le plus grand angle $-\alpha < \arg z < \alpha$ dans lequel la fonction $w = z^n$ reste injective vaut $\frac{2\pi}{n}$; il s'applique sur tout le plan des w dépourvu du demi-axe réel négatif.

La fonction inverse que l'on désigne, naturellement, par $z = \sqrt[n]{w}$ réalise une application bijective du domaine mentionné du plan des w sur l'intérieur de l'angle $-\frac{\pi}{n} < \arg z < \frac{\pi}{n}$ du plan des z .

Au point $z = 0$, la fonction $w = z^n$ est définie et analytique, mais l'application n'est pas conforme en ce point; en effet, la dérivée de la fonction z^n avec $n > 1$ est nulle pour $z = 0$.

10.57. Fonction exponentielle $w = e^z$ et son inverse. Ici on a

$$e^z = e^{x+iy} = e^x (\cos y + i \sin y).$$

En particulier, $|e^z| = e^x$, $\arg e^z = y + 2k\pi$ ($k=0, \pm 1, \pm 2, \dots$).

Toute droite horizontale $y = y_0$ du plan des z se transforme en la ligne $u = e^x \cos y_0$, $v = e^x \sin y_0$ dans le plan des w ; c'est une demi-droite issue de l'origine des coordonnées et faisant un angle y_0 avec le demi-axe positif des w . La bande définie par les inégalités $-\alpha < y < \alpha < \pi$ dans le plan des z se transforme biunivoquement

par la fonction $w = e^z$ sur l'angle

$$-\alpha < \arg w < \alpha$$

dans le plan des w . La plus grande bande $-\alpha < y < \alpha$ dans laquelle la fonction $w = e^z$ reste injective a $2\alpha = 2\pi$ pour largeur (de sorte que $\alpha = \pi$); la fonction $w = e^z$ réalise une application bijective de cette bande sur le domaine W que l'on obtient en enlevant au plan des w le demi-axe réel négatif. C'est dans ce domaine W que la fonction inverse désignée, naturellement, par $z = \ln w$ est définie et analytique. Pour les w réels positifs, cette fonction se confond avec la fonction bien connue $\ln w$ qu'on a définie dans 5.41. Ensuite, il résulte de la relation $|e^z| = e^x$ que $x = \ln |e^z|$; comme on a aussi $y = \arg e^z$, on voit que

$$\ln w = z = x + iy = \ln |e^z| + i \arg e^z = \ln |w| + i \arg w,$$

ce qui permet de dégager les parties réelle et imaginaire de la fonction $\ln w$. En inversant l'égalité $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$, on obtient la propriété correspondante du logarithme:

$$\ln(w_1 \cdot w_2) = \ln w_1 + \ln w_2.$$

Trouvons la dérivée du logarithme. Comme $(e^z)' = e^z$ (10.37), on a, d'après la formule 10.53 (2):

$$\ln(w)' = \frac{1}{(e^z)'} = \frac{1}{e^z} = \frac{1}{w}.$$

Compte tenu du théorème 10.25, il en résulte que la fonction $z = \ln w$ peut être mise, dans le domaine W , sous la forme de l'intégrale

$$\ln w = \ln w_0 + \int_{w_0}^w \frac{d\omega}{\omega},$$

où le chemin d'intégration est n'importe quelle ligne lisse par morceaux qui joint les points w_0 et w dans le domaine W . Si $w_0 = 1$, alors $\ln w_0 = 0$, et l'on a

$$\ln w = \int_1^w \frac{d\omega}{\omega}.$$

La dérivée de la fonction composée $\ln w(z)$ vaut $\frac{w'(z)}{w(z)}$, ce qui explique le terme « résidu logarithmique » de 10.44. La fonction $\ln w$ est maintenant définie pour tous les w sauf les nombres réels négatifs; les nombres négatifs n'ont pas encore de logarithmes. Un peu plus bas, nous indiquerons le domaine de définition complet du logarithme, y compris les nombres négatifs.

10.58. Fonction puissance générale $w = z^\lambda$, où λ est un nombre positif. Cette fonction est définie par la formule

$$w = e^{\lambda \ln z}$$

à condition que z ne soit pas un nombre réel négatif. La fonction $\ln z$ applique un angle $-\alpha < \arg z < \alpha < \min\left(\pi, \frac{\pi}{\lambda}\right)$ sur la bande $-\alpha < \operatorname{Im} \ln z < \alpha$; la multiplication par λ transforme cette bande en la bande $-\alpha\lambda < \operatorname{Im} \lambda \ln z < \alpha\lambda$, enfin la fonction exponentielle applique cette dernière sur l'angle

$$-\alpha\lambda < \arg w < \alpha\lambda$$

dans le plan des w . Le plus grand angle $-\alpha < \arg z < \alpha$ dans lequel la fonction $w = z^\lambda$ reste injective vaut $2\alpha = \frac{2\pi}{\lambda}$ ($\alpha = \frac{\pi}{\lambda}$) (pour $\lambda \geq 1$) et s'applique par cette fonction sur tout le plan des w dépourvu du demi-axe réel négatif.

En appliquant les règles usuelles, on trouve la dérivée de la fonction $w = z^\lambda$:

$$(e^{\lambda \ln z})' = e^{\lambda \ln z} (\lambda \ln z)' = \frac{\lambda}{z} e^{\lambda \ln z} = \lambda z^{\lambda-1}$$

10.59. Fonctions multivoques et surfaces riemanniennes.

a. Considérons la fonction

$$w = (z - z_0)^n, \quad n > 1,$$

dans un voisinage $|z - z_0| < \rho$ du point z_0 . Tout en étant injective dans l'angle $-\frac{\pi}{n} < \arg(z - z_0) < \frac{\pi}{n}$, cette fonction ne l'est plus dans tout le voisinage de z_0 , bien qu'elle y soit analytique. La circonférence $L = \{|z - z_0| = \rho\}$ s'applique sur la circonférence $\Lambda = \{|w| = \rho^n\}$ dans le plan des w , mais tandis que la variable z parcourt L une seule fois, la fonction $w = z^n$ parcourt la circonférence correspondant Λ n fois.

b. Une situation analogue a lieu pour un voisinage d'un zéro z_0 de multiplicité n de toute fonction analytique

$$w = f(z) = c_n (z - z_0)^n + c_{n+1} (z - z_0)^{n+1} + \dots, \quad c_n \neq 0, \quad n > 1.$$

En effet, si G est un voisinage du point z_0 dans lequel les fonctions $f(z)$ et $f'(z)$ n'ont pas de zéros autres que z_0 , et si $L \subset G$ est une courbe fermée contournant le point z_0 , alors, d'après 10.44, le résidu logarithmique de la fonction $f(z)$ par rapport au contour L vaut n :

$$\frac{1}{2\pi i} \int_L \frac{f'(\zeta)}{f(\zeta)} d\zeta = n.$$

Pour la raison de continuité, on a aussi, de même que dans 10.52 :

$$\frac{1}{2\pi i} \int \frac{f'(\zeta)}{f(\zeta) - w} d\zeta = n$$

si w est assez petit. Il en résulte d'après 10.44 que, compte tenu des multiplicités, la fonction $f(z)$ prend la valeur w dans le voisinage G n fois. Or, la fonction $f'(z)$ n'a qu'une seule racine z_0 dans le voisinage G . Donc la fonction $f(z)$ prend effectivement toute valeur $w \neq 0$ en n points différents. Dans de pareils cas, afin de rétablir la bijectivité de l'application, on introduit la notion de *surface riemannienne*. Sans formuler ici la définition générale, bornons-nous à deux exemples caractéristiques.

c. Pour la fonction $w = (z - z_0)^n$, la surface riemannienne est construite comme suit. On considère n exemplaires du plan des w coupés suivant le demi-axe réel négatif. On en identifie le point $w = 0$, puis l'on colle les feuillettes (c'est-à-dire qu'on effectue certaines identifications) de la manière suivante : le bord supérieur de la coupure du premier feuillet est collé au bord inférieur de la coupure du second, le bord supérieur de la coupure du second feuillet au bord inférieur de la coupure du troisième, etc., enfin le bord supérieur de la coupure du n -ième feuillet est collé au bord inférieur de la coupure du premier feuillet (fig. 10.9). (En réalité, il est impossible de réaliser un tel collage dans l'espace tridimensionnel, sans avoir des intersections superflues, mais cela n'a pas d'importance.) Ceci fait, le système initial des coordonnées est conservé sur chaque feuillet. On peut maintenant considérer l'application $w = z^n$ comme application bijective du plan achevé des z sur la surface riemannienne construite. Notamment, faisons correspondre au point $z = \rho > 0$ le point $w = \rho^n$ du premier feuillet. Lorsque z parcourt la circonférence $|z| = \rho$ dans le sens positif, $w = z^n$ parcourt une ligne sur le premier feuillet de la surface riemannienne toujours dans le sens positif, et au moment où le point z arrive à la frontière du domaine dans lequel $w = z^n$ est injective (10.56), c'est-à-dire que $\arg z$ prend la valeur $\frac{\pi}{n}$, le point correspondant w aura π pour argument, donc se trouvera sur la demi-droite le long de laquelle sont collés deux premiers feuillettes. Si le point z poursuit son parcours dans le même sens, le point w passe sur le second feuillet, puis sur le troisième, etc., et enfin, lorsque le point z fait un tour complet, w revient à sa position initiale sur le premier feuillet. Nous avons donc obtenu l'application bijective cherchée du plan des z sur la

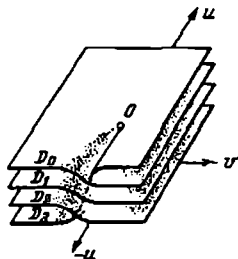


Fig. 10.9.

surface riemannienne que l'on a construite à la place du plan des w . Sur la surface riemannienne, on peut déjà définir la fonction inverse de $w = z^n$ notée $z = \sqrt[n]{w}$; elle applique d'une façon biunivoque la surface riemannienne sur le plan des z .

A chacun des n points différents de même coordonnée complexe $w \neq 0$ sur la surface riemannienne correspond l'un des n points différents $z_0, \dots, z_{n-1}$ du plan des z ; ces derniers ont le même module $\sqrt[n]{|w|}$, et leurs arguments sont donnés par l'égalité

$$\arg z_k = \frac{1}{n} (\arg w + 2k\pi), \quad k = 0, 1, \dots, n-1.$$

En particulier, pour $n = 2$ et $w = -1$, on a

$$z_{1,2} = \sqrt{-1} = i \sin \frac{(2k+1)\pi}{2} = \pm i$$

conformément à la règle fondamentale de l'algèbre des nombres complexes.

d. Afin de construire la surface riemannienne qui rend biunivoque la fonction $w = e^z$, on doit considérer un nombre infini (dénombrable)

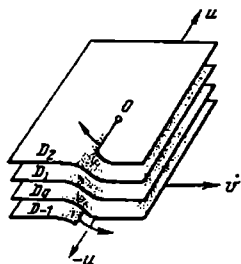


Fig. 10.10.

d'exemplaires du w -plan numérotés par tous les nombres entiers de $-\infty$ à $+\infty$ et coller ces feuillets au point $z = 0$ et suivant les coupures le long du demi-axe réel négatif: on colle le bord supérieur de la coupure du k -ième feuillet au bord inférieur de la coupure du $(k+1)$ -ième (fig. 10.10). Le domaine signalé dans 10.57 dans lequel la fonction $w = e^z$ est injective, c.-à-d. la bande $-\pi < y < \pi$, s'applique d'une façon biunivoque sur le feuillet numéro 0 dépourvu du demi-axe réel négatif. Si le point z se déplace le long de la verticale $x = x_0$ du point $y = 0$ à $y = \pi$, alors le point w parcourt, sur

le feuillet numéro 0, la demi-circonférence de rayon e^{x_0} jusqu'au bord supérieur de la coupure. Si le point z poursuit son déplacement vers le haut en passant par $y = \pi$, le point w traverse la coupure, arrive sur le second feuillet et décrit sur lui la circonférence de même rayon; lorsque le point z passe par $y = 3\pi$, le point w traverse encore une fois la coupure et se trouve sur le troisième feuillet, etc.

La fonction inverse $z = \ln w$ réalise une application bijective de la surface riemannienne construite sur le plan achevé des z . Aux valeurs de w sur le n -ième feuillet correspondent les valeurs de z dans la bande $(2n-1)\pi < y < (2n+1)\pi$. De la sorte, la fonction inverse de $w = e^z$ est une fonction à une infinité de déterminations, elle est désignée par $\text{Ln } w$; toutes ses valeurs pour un même w

sont données par la formule

$$\operatorname{Ln} w = \ln |w| + i \arg w,$$

où l'on tient compte de toutes les valeurs possibles de l'arg w .

En particulier, si w est un nombre réel négatif $w = -p$, $p > 0$, on a

$$\operatorname{Ln}(-p) = \ln |p| + (2k + 1)\pi i \quad (k = 0, \pm 1, \pm 2, \dots),$$

et il n'y a aucune raison tant soit peu naturelle de préférer l'une de ces valeurs à toutes les autres.

e. Les surfaces riemanniennes pour d'autres fonctions analytiques sont construites de manière analogue. Dans le cas général, pour une application $w = f(z)$, on construit les surfaces riemanniennes pour la variable z et pour la fonction w .

Afin de construire la surface riemannienne de la variable z , on prend un nombre suffisant de feuilletts identiques qui correspondent au domaine de définition de la fonction $f(z)$, on y fait les coupures partageant ce domaine en parties dans chacune desquelles la fonction $f(z)$ est injective; on doit ensuite coller les bords de ces coupures (cf. exemples c et d) dans le but de garantir la bijectivité de l'application de toute la surface riemannienne de la variable z sur toute la surface riemannienne de la variable w .

Si, en contournant un point dans le plan des z suivant une circonférence de rayon arbitrairement petit, les valeurs de la fonction $f(z)$ passent sur un autre feuillet, alors ce point est appelé *point de branchement* de la fonction $f(z)$. Les points de branchement jouent un rôle particulier dans la structure d'une surface riemannienne. Pour tout point z_0 qui n'est pas un point de branchement de la fonction $f(z)$, il existe un voisinage de ce point dans lequel la fonction $f(z)$ se réduit à une famille de fonctions analytiques; un point de branchement n'a pas de voisinage possédant cette propriété.

Dans les exemples c et d les points de branchement étaient 0 et ∞ .

La structure des surfaces riemanniennes pour de différentes fonctions est étudiée plus en détail dans [3]. Pour l'état actuel de la théorie générale des surfaces riemanniennes voir [22].

Exercices

1. Si le rayon de convergence de la série $f(z) = \sum_0^{\infty} a_n z^n$ vaut 1 et que tous les a_n soient positifs, alors $z = 1$ est un point singulier.

2. Si la série $\sum_0^{\infty} a_n z^n$ a 1 pour rayon de convergence et le seul point singulier sur la circonférence $|z| = 1$ est un pôle d'ordre m , alors $|a_n| \leq A n^{m-1}$, où A est une constante.

3. (*Principe du maximum.*) Soit $f(z)$ une fonction analytique dans le domaine G limité par un contour Γ , et soit

$$M = \sup_{z \in \Gamma} |f(z)|.$$

Alors $|f(z)| \leq M$ dans le domaine G . Si $|f(z_0)| = M$ en un point intérieur $z_0 \in G$, alors $f(z)$ est constante dans le domaine G .

4. Si une fonction $f(z)$ qui est analytique à l'intérieur du cercle $|z| \leq R$ y satisfait à l'inégalité $|f(z)| \leq M$, et si $f(0) = 0$, alors

$$|f(z)| < \frac{M}{R} |z|$$

(H. Schwarz).

5. Si une fonction $f(z)$ est analytique dans le cercle $|z| < R$, alors la fonction $M(\rho)$, où $M(\rho) = \max_{|z|=\rho} |f(z)|$, est une fonction convexe de $\ln \rho$

(J. Hadamard).

6. Une fonction réelle $u(x, y)$ est dite différentiable en un point (x_0, y_0) s'il existe des nombres A_u et B_u tels que, dans un voisinage du point x_0, y_0 , on a la représentation

$$u(x_0 + h, y_0 + k) = u(x_0, y_0) + A_u h + B_u k + \varepsilon(h, k),$$

où $\frac{\varepsilon(h, k)}{|h| + |k|} \rightarrow 0$ pour $h \rightarrow 0, k \rightarrow 0$. Démontrer que, pour que la fonction $w = u(x, y) + iv(x, y)$ ait la dérivée $w'(z_0)$, il faut et il suffit que $u(x, y)$ et $v(x, y)$ soient différentiables au point $z_0 = (x_0, y_0)$ et que l'on ait $A_u = B_v, A_v = -B_u$.

7. Démontrer qu'une fonction homographique $w = \frac{az+b}{cz+d}$ transforme la famille de toutes les circonférences (y compris les droites que l'on peut considérer comme circonférences centrées à l'infini) en elle-même.

8. La croissance d'une fonction $f(z) = \sum_0^{\infty} a_n z^n$ peut être évaluée par les fonctions

$$M(\rho) = \max_{|z|=\rho} |f(z)|,$$

$$M_1(\rho) = \sum_0^{\infty} |a_n| \cdot |z|^n.$$

Evidemment, $M(\rho) \leq M_1(\rho)$. Mais $M_1(\rho)$ ne croît pas trop vite relativement à $M(\rho)$: pour tout $\delta > 0$, on a

$$M_1(r) \leq \frac{r+\delta}{\delta} M(r+\delta).$$

9. Supposons qu'une fonction $f(z)$ soit analytique dans le domaine G limité par un contour L , à l'exception éventuelle d'un nombre fini de pôles, et qu'elle soit non nulle sur L . Lorsque z parcourt L dans le sens positif, le point $w = f(z)$ décrit une courbe fermée. Le nombre de tours complets que fait cette courbe (autour du point $w = 0$) est égal au nombre des zéros de la fonction $f(z)$ dans le domaine G moins le nombre des pôles (*principe de l'argument*).

10. On sait qu'une suite de fonctions analytiques $f_n(z)$ ($n = 1, 2, \dots$) converge vers une fonction $f(z) \neq 0$ uniformément à l'intérieur d'un domaine G (10.36. a). Soit Z l'ensemble de tous les zéros de toutes les fonctions $f_n(z)$ dans le domaine G . Alors l'ensemble des zéros de la fonction $f(z)$ (tout zéro de multi-

plicité m est compté m fois) dans G se confond avec l'ensemble des points limites de Z à l'intérieur de G .

11. Si la série $\sum_0^{\infty} g^{(n)}(z)$ converge pour au moins un point z d'analyticité de la fonction $g(z)$, alors $g(z)$ est une fonction entière, et la série mentionnée converge pour tout z uniformément dans tout cercle (G. Polya).

12. Pour les fonctions $w = \frac{1}{2} \left(z + \frac{1}{z} \right)$, $w = \cos z$ et leurs inverses $z = w \pm \sqrt{w^2 - 1}$, $z = \operatorname{Arccos} w$, trouver les domaines dans lesquels elles sont injectives, construire les surfaces riemanniennes et indiquer les points de branchement.

13. Démontrer l'égalité

$$i \operatorname{Arccos} w = \operatorname{Ln} (w \pm \sqrt{w^2 - 1}).$$

14. Si les fonctions $f_1(z), \dots, f_n(z)$ sont analytiques dans un domaine G , alors la fonction $|f_1(z)| + \dots + |f_n(z)|$ n'atteint pas son maximum à l'intérieur de G .

15. Une fonction analytique $f(z)$ définie dans tout le plan des z applique le plan tout entier dans le demi-plan supérieur. Démontrer que $f(z)$ est une constante.

16. Dédire le théorème sur l'existence d'une racine d'un polynôme algébrique $P(z)$ (5.86) en partant du théorème de Liouville 10.38.

17. Dédire le théorème 5.86 en partant du théorème sur le résidu logarithmique (10.44).

18. Dédire le théorème 5.86 en partant du principe du maximum (exercice 3).

19. Une fonction analytique $f(z)$ est définie dans tout le plan à l'exception des points $z_1, z_2, \dots, z_n, \dots$ où elle a les pôles simples de résidus $b_1, \dots, b_n, \dots$ respectivement. Supposons que l'ensemble $\{z: |f(z)| \leq C\}$, pour un certain C , contienne une famille de circonférences Γ_n de centre $z = 0$ et de rayon $R_n \rightarrow \infty$. Démontrer que

$$f(z) = f(0) + \sum_{n=1}^{\infty} b_n \left\{ \frac{1}{z - z_n} + \frac{1}{z_n} \right\}$$

(décomposition de $f(z)$ en éléments simples).

20. Supposons qu'une fonction $g(z)$ soit analytique dans tout le plan et possède les zéros simples aux points $z_1, z_2, \dots$. Supposons, de plus, que la fonction

$f(z) = \frac{g'(z)}{g(z)}$ satisfasse à l'hypothèse de l'exercice 19. Montrer que

$$g(z) = g(0) e^{\frac{g'(0)}{g(0)} z} \prod_{n=1}^{\infty} \left\{ \left(1 - \frac{z}{z_n} \right) e^{\frac{z}{z_n}} \right\}.$$

21. Etablir les décompositions suivantes :

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{\pi^2 n^2} \right),$$

$$\cos z = \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{\pi^2 \left(n + \frac{1}{2} \right)^2} \right).$$

22. Supposons qu'une fonction $f(z)$ soit analytique dans le domaine $G = \left\{ z : -\alpha \leq \arg z \leq \alpha < \frac{\pi}{2} \right\}$ et vérifie les conditions :

$$|f(z)| < Ae^{B|z|} \text{ dans } G;$$

$$|f(z)| \leq 1 \text{ sur les demi-droites } \arg z = \pm \alpha.$$

Alors $|f(z)| \leq 1$ dans tout le domaine G (Phragmén-Lindelöf).

23. Supposons qu'une fonction $f(z)$ soit analytique dans le demi-plan supérieur G et vérifie les conditions :

$$|f(x)| \leq c \quad (-\infty < x < \infty)$$

et, quel que soit $\varepsilon > 0$,

$$|f(z)| \leq Ae^{\varepsilon|z|} \text{ pour tous les } z \in G.$$

Alors $|f(z)| \leq c$ pour tous les $z \in G$.

24. Soit $f(z)$ une fonction analytique entière vérifiant les conditions :

$$|f(z)| \leq e, \quad \inf |f(x)| = 0 \quad (-\infty < x < \infty)$$

et, quel que soit $\varepsilon > 0$,

$$|f(z)| \leq Ae^{\varepsilon|z|} \text{ pour tous les } z \in G.$$

Alors $f(z) = 0$.

Historique

À la fin du XVII^e siècle et au début du XVIII^e, on possède déjà beaucoup de faits concernant les propriétés des fonctions élémentaires dans lesquelles la variable réelle est remplacée par la variable imaginaire (complexe). Faute de représentation géométrique d'un nombre complexe (qui n'est indiquée par Gauss qu'en 1799), les mathématiciens de l'époque raisonnent par analogie et aboutissent souvent à des paradoxes. Ainsi, Leibniz et Jean Bernoulli disputent entre eux si les logarithmes des nombres négatifs sont imaginaires ou réels, et cela sans posséder aucune définition du logarithme dans le domaine complexe (une telle définition prenant en considération la multivocité du logarithme est donnée par Euler en 1749). Les « conditions de Cauchy-Riemann » pour les fonctions représentées par des séries de puissances sont trouvées par d'Alembert (1752) et Euler (1777). Le terme « fonction analytique » est introduit par Lagrange dans sa « Théorie des fonctions analytiques contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits ou d'évanouissants, de limites ou de fluxions, et réduits à l'analyse algébrique des quantités finies » (1797), où l'on sous-entend par fonction analytique la somme d'une série de puissances. D'ailleurs, Lagrange ne pose même pas la question de convergence de ses séries (sinon, il serait obligé de revenir à la notion de limite qu'il a chassée). Les travaux de Cauchy (à partir de 1814) contenant plusieurs théorèmes fondamentaux de la théorie des fonctions analytiques sont basés sur la même définition, mais l'auteur met à nu le sens de la convergence de séries et la signification géométrique de la variable complexe; c'est Cauchy qui considère, pour la première fois, l'intégrale curviligne par rapport à la variable complexe, dont il réduit la définition à celle de l'intégrale usuelle par rapport à la variable réelle en séparant les parties réelle et imaginaire. Il indique la méthode de calcul des intégrales de fonctions analytiques le long de contours fermés à l'aide de résidus. D'autre part, Cauchy découvre la relation entre l'analyticité d'une fonction et sa dérivabilité par rapport à la variable complexe et appelle monogénéité cette dernière propriété. L'analyse des points singuliers d'une fonction univoque à l'aide de sa série de Laurent (1843) est réalisée simultanément par le mathématicien russe Sokhotski et le mathématicien italien Casorati (1868) et, un peu plus tard, par Weierstrass (1876).

Intégrales impropres

Le calcul intégral ne vaudrait pas grand'chose si, au lieu d'être cherchée sous forme finie, la primitive demandait toujours un passage à la limite; il n'y aurait alors aucun progrès... Le mérite immense de Newton et Leibniz est de rendre possible la représentation explicite du résultat d'une opération qui est, en fait, irréalisable... 250 années passèrent depuis, et pas mal d'hommes forts refirent le chemin de Newton et Leibniz. Personne ne put continuer l'œuvre des fondateurs de l'analyse infinitésimale, ni indiquer une autre méthode pour mettre sous forme finie le résultat de la sommation d'infiniment petits... Seul Cauchy, le plus grand analyste du XIX^e siècle, eut l'honneur de donner un procédé original, qu'il nomma calcul des résidus, dont on peut se servir pour mettre sous forme finie le résultat de la sommation d'un nombre infiniment grand de termes infiniment petits, et cela dans une quantité de cas où la méthode éprouvée de primitives de Newton et Leibniz est en défaut.

N. Lousine

Isaac Newton, mathématicien et naturaliste (1643)

§ 11.1. Intégrales impropres de première espèce

11.11. Soit $y = f(x)$ une fonction définie pour $a \leq x \leq b < \infty$ et prenant des valeurs complexes. La définition de l'intégrale d'une fonction à valeurs complexes est connue depuis 10.21; notamment, si $f(x) = u(x) + i v(x)$, où $u(x)$ et $v(x)$ sont réelles, on pose par définition

$$\int_a^b f(x) dx = \int_a^b u(x) dx + i \int_a^b v(x) dx.$$

L'intégrabilité de la fonction $f(x)$ sur l'intervalle $[a, b]$ est donc équivalente à l'intégrabilité, sur cet intervalle, de deux fonctions $u(x)$ et $v(x)$.

Supposons qu'une fonction $f(x)$ soit intégrable (par exemple, continue par morceaux) sur tout intervalle fini $a \leq x \leq b < \infty$, a étant fixe, b arbitraire. Nous tenons à attribuer un sens à l'*intégrale impropre de première espèce* »

$$\int_a^{\infty} f(x) dx. \quad (1)$$

Considérons la fonction complexe de la variable $X \geq a$:

$$I(X) = \int_a^X f(x) dx. \quad (2)$$

Définition. Si, pour $X \rightarrow \infty$, la fonction $I(X)$ tend vers une limite finie $I \in \mathbb{C}$, on dit que l'intégrale impropre (1) *converge* et l'on pose par définition

$$\int_a^{\infty} f(x) dx = \lim_{X \rightarrow \infty} I(X) = \lim_{X \rightarrow \infty} \int_a^X f(x) dx = I. \quad (3)$$

Si, pour $X \rightarrow \infty$, la fonction $I(X)$ n'a pas de limite finie, on dit que l'intégrale (1) *diverge*, et l'on ne lui attribue aucune valeur.

Exemples. a. Intégrale impropre $\int_1^{\infty} x^{\alpha} dx$.

Vu que

$$\int_1^X x^{\alpha} dx = \begin{cases} \frac{X^{\alpha+1}-1}{\alpha+1} & \text{pour } \alpha \neq -1, \\ \ln X & \text{pour } \alpha = -1, \end{cases}$$

l'intégrale converge, si $\alpha < -1$, et diverge, si $\alpha \geq -1$.

b. Intégrale impropre $\int_0^{\infty} \cos x dx$.

Comme

$$\int_0^X \cos x dx = \sin X$$

n'a pas de limite pour $X \rightarrow \infty$, l'intégrale diverge.

c. Intégrale impropre $\int_0^{\infty} e^{ix} dx$.

Comme

$$\int_0^X e^{ix} dx = \frac{e^{iX}-1}{i}$$

n'a pas de limite pour $X \rightarrow \infty$, l'intégrale diverge.

11.12. Critère de Cauchy et conditions équivalentes de convergence. *Les propositions suivantes sont équivalentes :*

a. Il existe un I tel que, pour tout $\varepsilon > 0$, on peut trouver un $X_0 > a$ tel que, quel que soit $X \geq X_0$, on a

$$|I - I(X)| = \left| I - \int_a^X f(x) dx \right| < \varepsilon.$$

b. Pour toute suite $X_n \rightarrow \infty$, les nombres $I(X_n)$ tendent vers une limite (finie).

c. Pour toute suite $X_n \rightarrow \infty$, la série

$$\sum_n [I(X_{n+1}) - I(X_n)] = \sum_n \int_{X_n}^{X_{n+1}} f(x) dx$$

converge.

d. Quel que soit un $\varepsilon > 0$, il existe un $X > a$ tel que, pour $X' \geq X$, $X'' \geq X$, on a

$$|I(X') - I(X'')| = \left| \int_{X'}^{X''} f(x) dx \right| < \varepsilon$$

(critère de Cauchy).

La proposition a n'est rien d'autre que la définition de la limite de la fonction $I(X)$ à l'infini. La proposition d est le critère de Cauchy d'existence de cette limite. La proposition b représente la condition équivalente en termes des suites (4.66). Enfin, la proposition c exprime la relation usuelle entre la convergence d'une série et celle de la suite de ses sommes partielles. Les propositions a-d sont donc équivalentes.

Il importe, dans la proposition b, qu'il s'agit de toute suite $X_n \rightarrow \infty$. Si l'on sait que, pour une certaine suite $X_n \rightarrow \infty$, les nombres $I(X_n)$ tendent vers une limite, cela ne signifie pas encore que l'intégrale (1) converge. Par exemple,

$$\int_0^{2\pi n} \cos x dx = 0,$$

mais l'intégrale 11.11.b diverge.

11.13. Cas de fonction à intégrer non négative.

a. Soit une intégrale impropre

$$\int_a^\infty f(x) dx, \quad (1)$$

où la fonction $f(x)$ est non négative. Alors la primitive $I(X) = \int_a^X f(x) dx$ ne décroît pas; si $I(X)$ n'est pas bornée sur (a, ∞) , alors sa limite pour $X \rightarrow \infty$ est $+\infty$, et nous disons que l'intégrale (1) *diverge vers $+\infty$* ; si $I(X)$ est bornée sur (a, ∞) , alors $\sup I(X) = \lim_{X \rightarrow \infty} I(X)$ (4.53), et l'intégrale (1) converge.

b. Le critère de comparaison est formulé de la manière suivante: si deux fonctions $f_1(x)$ et $f_2(x)$ non négatives, intégrables sur tout intervalle fini et telles que $f_1(x) \leq c f_2(x)$ pour $x \geq X_0$ sont données sur (a, ∞) , alors la convergence de l'intégrale de $f_2(x)$ implique la convergence de l'intégrale de $f_1(x)$, et la divergence de l'intégrale de $f_1(x)$ implique la divergence de l'intégrale de $f_2(x)$.

Toutes ces conclusions résultent de l'inégalité

$$I_1(X) - I_1(X_0) = \int_{X_0}^X f_1(x) dx \leq c \int_{X_0}^X f_2(x) dx = c(I_2(X) - I_2(X_0)),$$

qui a lieu pour tout $X > X_0$, ainsi que de la convergence d'une intégrale avec la primitive $I(X)$ bornée (a).

c. Exemple. L'intégrale d'une fonction rationnelle

$$\int_a^\infty \frac{P(x)}{Q(x)} dx \quad (a > 0), \quad (2)$$

où $P(x)$ et $Q(x) \neq 0$ sont des polynômes (à coefficients complexes) de degré p et q respectivement, converge pour $q \geq p + 2$ et diverge pour $q \leq p + 1$.

En effet, si $q \geq p + 2$, alors $\left| \frac{P(x)}{Q(x)} \right| \leq \frac{c}{x^2}$, et l'intégrale (2) converge, car on peut choisir à titre de comparaison la fonction $\frac{1}{x^2}$ (11.11.a). Soit maintenant $q \leq p + 1$. Alors la fonction rationnelle

$x \frac{P(x)}{Q(x)}$ a une partie entière non nulle; on peut poser

$$x \frac{P(x)}{Q(x)} = a_0 x^m + a_1 x^{m-1} + \dots + a_{m-1} x + a_m + \frac{P_1(x)}{Q_1(x)},$$

où $m \geq 0$, $a_0 \neq 0$, et le degré de $P_1(x)$ est inférieur à celui de $Q_1(x)$. Il en résulte

$$\begin{aligned} \frac{P(x)}{Q(x)} &= a_0 x^{m-1} + \dots + a_{m-1} + \frac{a_m}{x} + \frac{1}{x} \frac{P_1(x)}{Q_1(x)}, \\ \int_a^X \frac{P(x)}{Q(x)} dx &= \frac{a_0}{m} (X^m - a^m) + \dots + a_{m-1} (X - a) + a_m \ln \frac{X}{a} + \\ &\quad + \int_a^X \frac{1}{x} \frac{P_1(x)}{Q_1(x)} dx. \end{aligned}$$

Le dernier terme, comme on l'a vu plus haut, a une limite finie (le degré du dénominateur dépasse au moins de deux unités celui du numérateur), la somme des autres termes n'a pas de limite, son module tendant vers l'infini. Par conséquent, l'intégrale (2) diverge.

d. E x e m p l e. Pour un polynôme $P(x) = a_0x^m + \dots + a_m$ à coefficients réels et avec $a_0 > 0$, l'estimation suivante a lieu si x sont grands :

$$\frac{c_1}{x^{m/p}} \leq \frac{1}{\sqrt[p]{a_0x^m + \dots + a_m}} \leq \frac{c_2}{x^{m/p}},$$

donc l'intégrale

$$\int_a^\infty \frac{dx}{\sqrt[p]{P(x)}}$$

converge pour $\frac{m}{p} > 1$ et diverge pour $\frac{m}{p} \leq 1$.

11.14. Critère intégral de convergence d'une série numérique. Soit $a_1 + a_2 + \dots$ une série à termes positifs non

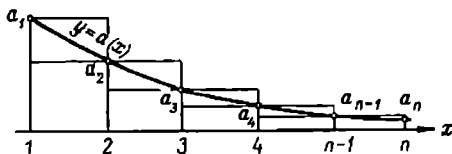


Fig. 11.1

croissants, de sorte que $a_{n+1} \leq a_n$ ($n=1, 2, \dots$) et soit $y=a(x)$ une fonction positive non croissante et telle que $a(n)=a_n$. La fig. 11.1 montre que

$$a_2 + \dots + a_n \leq \int_1^n a(x) dx \leq a_1 + a_2 + \dots + a_{n-1}. \quad (1)$$

Il résulte de (1): si l'intégrale $\int_1^\infty a(x) dx$ converge, il en est de même de la série $\sum_1^\infty a_n$, et si l'intégrale $\int_1^\infty a(x) dx$ diverge, il en est de même de la série $\sum_1^\infty a_n$. C'est le critère intégral de Cauchy de convergence d'une série numérique.

Exemples. a. On sait que $\int_1^{\infty} \frac{dx}{x^\alpha}$ converge pour $\alpha > 1$ et diverge pour $\alpha \leq 1$ (11.11.a). On retrouve donc le résultat sur la convergence de la série $\sum_1^{\infty} \frac{1}{n^\alpha}$ pour $\alpha > 1$ et sa divergence pour $\alpha \leq 1$ (6.15.a).

b. La relation

$$\int_a^{\infty} \frac{dx}{x \ln^\alpha x} = \int_{\ln a}^{\infty} \frac{du}{u^\alpha} \quad (u = \ln x, \alpha > 1)$$

montre que l'intégrale impropre $\int_a^{\infty} \frac{dx}{x \ln^\alpha x}$ converge pour $\alpha > 1$ et diverge pour $\alpha \leq 1$. La série $\sum_2^{\infty} \frac{1}{n \ln^\alpha n}$ converge donc pour $\alpha > 1$ et diverge pour $\alpha \leq 1$ (6.15.c).

c. En partant des considérations géométriques du même type qu'au début du présent numéro, on peut obtenir de différentes

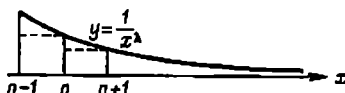


Fig. 11.2.

estimations des restes de séries convergentes et des sommes partielles de séries convergentes ou non. Par exemple, on a directement de la fig. 11.2:

$$\sum_{k=n}^{\infty} \frac{1}{k^\lambda} \leq \int_{n-1}^{\infty} \frac{dx}{x^\lambda} = \frac{1}{\lambda-1} \left[\frac{1}{(n-1)^{\lambda-1}} - 1 \right] \quad (\lambda > 1).$$

D'après la fig. 11.1, on a, pour $a(x) = \frac{1}{x}$:

$$\sum_{k=2}^n \frac{1}{k} \leq \int_1^n \frac{dx}{x} = \ln n \leq \sum_{k=1}^{n-1} \frac{1}{k}.$$

De la même manière, il résulte de la fig. 11.3:

$$\sum_{k=1}^n \left\{ a_k - \int_k^{k+1} a(x) dx \right\} \leq a_1,$$

le premier membre étant croissant avec n . En désignant sa limite par C et en sommant on voit que

$$\sum_{k=1}^n a_k = \int_1^{n+1} a(x) dx + C - \gamma_n, \quad \gamma_n \searrow 0. \quad (2)$$

Pour la série $a_k = \frac{1}{k}$, la constante C fut calculée pour la première fois par Euler (1734): $C = 0,5772 \dots$. La constante d'Euler se ren-

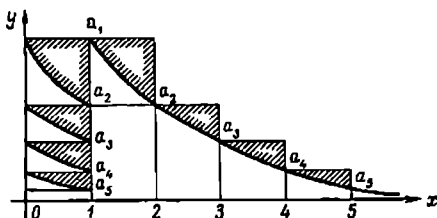


Fig. 11.3.

contre souvent en analyse. L'égalité (2) prend la forme

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \ln(n+1) + 0,5772 \dots - \gamma_n, \quad \gamma_n \searrow 0.$$

11.15. Convergence absolue et semi-convergence d'intégrales impropres. Soit une intégrale impropre

$$\int_a^\infty f(x) dx, \quad (1)$$

$f(x)$ étant une fonction complexe continue par morceaux. Si l'intégrale

$$\int_a^\infty |f(x)| dx \quad (2)$$

converge, alors l'intégrale (1) converge également, puisque, quels que soient X' et X'' , on a d'après 10.21 (5):

$$\left| \int_{X'}^{X''} f(x) dx \right| \leq \int_{X'}^{X''} |f(x)| dx$$

et l'on peut appliquer le critère de Cauchy. Dans ce cas, l'intégrale (1) est dite *absolument convergente* et la fonction $f(x)$ *absolument intégrable* sur (a, ∞) . Pourtant, l'intégrale (1) peut bien converger

sans que l'intégrale (2) converge; dans ce cas, l'intégrale (1) est dite *semi-convergente*. Il existe des critères de convergence applicables à des intégrales semi-convergentes.

11.16. Critère de Leibniz. Soit une intégrale impropre

$$\int_a^{\infty} f(x) dx, \quad (1)$$

la fonction $f(x)$ étant réelle. Supposons que la suite $a_1 < a_2 < \dots < a_n < \dots$ de toutes les racines de la fonction $f(x)$ sur (a, ∞)

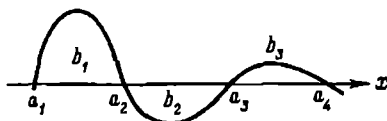


Fig. 11.4.

soit telle que $a_n \rightarrow \infty$, $f(x) > 0$ pour $a_{2n-1} < x < a_{2n}$ et $f(x) < 0$ pour $a_{2n} < x < a_{2n+1}$ (fig. 11.4). Les nombres

$$b_n = \int_{a_n}^{a_{n+1}} f(x) dx$$

forment une suite alternée. S'ils satisfont à l'hypothèse du critère de Leibniz pour une série numérique (6.23), c.-à-d. si $|b_n| \searrow 0$ et $|b_n| \geq |b_{n+1}|$ pour tous les $n = N, N+1, N+2, \dots$, alors l'intégrale (1) converge.

Pour démontrer la proposition, choisissons, pour un X donné, un n tel que $a_n \leq X \leq a_{n+1}$. Alors

$$\int_a^X f(x) dx = \int_a^{a_1} f(x) dx + [b_1 + \dots + b_{n-1}] + \int_{a_n}^X f(x) dx.$$

Le premier terme du second membre est constant. La somme entre crochets a une limite, lorsque $n \rightarrow \infty$, en vertu du critère de Leibniz pour les séries. Le dernier terme ne dépasse pas en module la quantité

$$\int_{a_n}^{a_{n+1}} |f(x)| dx = |b_n|,$$

donc, par hypothèse, tend vers zéro pour $n \rightarrow \infty$. Par conséquent, la limite du premier membre existe, ce qu'il nous fallait.

Exemples. a. Considérons l'intégrale

$$\int_a^{\infty} g(x) \sin x \, dx, \quad (2)$$

où $g(x) > 0$ et $g(x) \searrow 0$ pour $x \rightarrow \infty$. Les racines successives de la fonction $g(x) \sin x$ pour $x > a$ sont $x = k\pi$, $(k+1)\pi$, ... On a, évidemment,

$$g(x) |\sin x| \geq g(x+\pi) |\sin(x+\pi)|,$$

d'où

$$\begin{aligned} |b_m| &= \int_{m\pi}^{(m+1)\pi} g(x) |\sin x| \, dx \geq \int_{m\pi}^{(m+1)\pi} g(x+\pi) |\sin(x+\pi)| \, dx = \\ &= \int_{(m+1)\pi}^{(m+2)\pi} g(x) |\sin x| \, dx = |b_{m+1}|; \end{aligned}$$

ensuite, on a

$$|b_m| = \int_{m\pi}^{(m+1)\pi} g(x) |\sin x| \, dx \leq \pi g(m\pi) \rightarrow 0 \text{ pour } m \rightarrow \infty.$$

L'hypothèse du critère de Leibniz est donc satisfaite, et l'intégrale (2) converge.

Si $\int_a^{\infty} g(x) \, dx < \infty$, alors l'intégrale (2) converge absolument.

Montrons que si $\int_a^{\infty} g(x) \, dx = \infty$, alors l'intégrale

$$\int_a^{\infty} g(x) |\sin x| \, dx \quad (3)$$

diverge. Dans l'intervalle $(m-1)\pi + \frac{\pi}{6} \leq x \leq m\pi - \frac{\pi}{6}$ de longueur $\frac{2}{3}\pi$, on a les inégalités $|\sin x| \geq \frac{1}{2}$, $g(x) \geq g(m\pi)$; par conséquent

$$\begin{aligned} \int_{(m-1)\pi}^{m\pi} g(x) |\sin x| \, dx &\geq \int_{(m-1)\pi + \frac{\pi}{6}}^{m\pi - \frac{\pi}{6}} g(x) |\sin x| \, dx \geq \\ &\geq g(m\pi) \cdot \frac{1}{2} \cdot \frac{2}{3} \pi. \end{aligned}$$

Or, l'intégrale

$$\int_a^{\infty} g(x\pi) dx = \frac{1}{\pi} \int_{a\pi}^{\infty} g(u) du \quad (u = \pi x)$$

étant divergente, la série des termes $g(m\pi)$ diverge. Donc, d'après 11.12.c, l'intégrale (3) diverge elle aussi.

b. L'intégrale

$$\int_a^{\infty} \sin(x^\gamma) dx, \quad \gamma > 0, \quad (4)$$

se réduit par la substitution $x^\gamma = u$ à l'intégrale

$$\frac{1}{\gamma} \int_b^{\infty} \sin u \cdot u^{\frac{1}{\gamma}-1} du$$

qui converge, on l'a vu, pour $\frac{1}{\gamma} - 1 < 0$, c.-à-d. pour $\gamma > 1$ (on peut aussi appliquer le critère de Leibniz directement à l'intégrale (4)).

Notons que dans l'intégrale convergente (4), la fonction $\sin(x^\gamma)$ ne tend pas vers zéro pour $x \rightarrow \infty$, ce que l'analogie formelle avec les séries nous pourrait faire attendre. En réalité, l'analogue du théorème: « le terme général d'une série numérique convergente tend vers zéro », est non pas la proposition: « la fonction à intégrer d'une intégrale impropre convergente tend vers zéro pour $x \rightarrow \infty$ » (ce qui est faux comme on l'a vu ci-dessus), mais la proposition: « étant donnée une intégrale impropre convergente, l'intégrale de sa fonction sous le signe somme étendue sur un intervalle de longueur fixe tend vers zéro lorsque l'intervalle s'éloigne à l'infini ». En d'autres termes, pour tout $h > 0$, on a

$$\lim_{x \rightarrow \infty} \int_x^{x+h} f(u) du = 0,$$

ce qui résulte immédiatement du critère de Cauchy 11.12.d.

11.17.a. Critère d'Abel-Dirichlet. Soit intégrale impropre

$$\int_a^{\infty} g(x) s(x) dx. \quad (1)$$

Si la fonction (complexe) $g(x)$ possède une dérivée continue par morceaux absolument intégrable sur (a, ∞) et tend vers 0 pour $x \rightarrow \infty$, et si $s(x)$ a une primitive bornée $G(x)$, $|G(x)| \leq C$, alors l'intégrale (1) converge.

En intégrant par parties on a :

$$\int_p^q g(x) s(x) dx = g(x) G(x) \Big|_p^q - \int_p^q G(x) g'(x) dx \quad (p < q).$$

Le premier terme à droite admet la majoration

$$|g(q)G(q) - g(p)G(p)| \leq 2C \max\{|g(p)|, |g(q)|\},$$

donc tend vers zéro pour $p, q \rightarrow \infty$. Le terme intégral admet la majoration suivante :

$$\left| \int_p^q G(x) g'(x) dx \right| \leq C \int_p^q |g'(x)| dx.$$

Par conséquent, il tend également vers zéro pour $p \rightarrow \infty, q \rightarrow \infty$. Ainsi, l'intégrale (1) satisfait à la condition de Cauchy 11.12.d, donc converge.

b. Exemple. Considérons l'intégrale

$$\int_a^\infty \frac{P(x)}{Q(x)} e^{i\alpha x} dx \quad (\alpha \in \mathbb{R}, \alpha \neq 0), \quad (2)$$

où $P(x)$ et $Q(x)$ sont des polynômes en x (à coefficients complexes) tels que le degré du polynôme $Q(x)$ dépasse d'une unité celui de $P(x)$ et que $Q(x)$ n'a pas de racines dans l'intervalle (a, ∞) .

Posons $g(x) = \frac{P(x)}{Q(x)}$, $s(x) = e^{i\alpha x}$. Alors $g'(x) = \frac{P'Q - PQ'}{Q^2}$ est une fonction rationnelle dont le degré du dénominateur dépasse au moins de deux unités celui du numérateur. C'est pourquoi l'intégrale de $g'(x)$ est absolument convergente (11.13.c). La fonction $s(x) = e^{i\alpha x}$ a la primitive bornée $\frac{e^{i\alpha x}}{i\alpha}$. L'hypothèse du critère n'est donc satisfaite, et l'intégrale (2) converge.

c. Si la fonction $g(x)$ est réelle, on peut modifier l'hypothèse du critère a ; notamment, il suffit de supposer que $g(x)$ tende vers zéro d'une façon monotone et possède une dérivée continue par morceaux. En effet, dans l'hypothèse formulée, la fonction $g'(x)$ ne change pas de signe, et l'expression

$$\int_a^X |g'(x)| dx = \left| \int_a^X g'(x) dx \right| = |g(X) - g(a)|$$

a une limite finie pour $X \rightarrow \infty$, de sorte que $|g'(x)|$ est intégrable

Les conditions à imposer à la fonction $s(x)$ restent les mêmes

11.18. Intégration par parties et par substitution dans une intégrale impropre En fait, nous avons déjà utilisé plus d'une fois les procédés mentionnés dans des cas concrets; à présent, faisons quelques remarques d'ordre général. Il suffit de prouver qu'il soit possible d'effectuer les opérations en question pour un intervalle fini (a, X) et de passer ensuite à la limite pour $X \rightarrow \infty$. Ainsi, si l'intégration par parties sur l'intervalle (a, X) donne

$$\int_a^X f(x) g(x) dx = f(x) G(x) \Big|_a^X - \int_a^X f'(x) G(x) dx,$$

on peut noter

$$\int_a^\infty f(x) g(x) dx = f(x) G(x) \Big|_a^\infty - \int_a^\infty f'(x) G(x) dx$$

dans la supposition qu'au moins deux limites sur trois existent (alors la troisième existe évidemment elle aussi). Le premier terme à droite est défini comme suit:

$$f(x) G(x) \Big|_a^\infty = \lim_{X \rightarrow \infty} f(x) G(x) \Big|_a^X = \lim_{X \rightarrow \infty} [f(X) G(X) - f(a) G(a)]$$

D'une façon analogue, si, après une substitution $x = x(u)$ nous aboutissons à l'égalité

$$\int_a^X f(x) dx = \int_b^U f(x(u)) x'(u) du$$

et que, de plus, $X \rightarrow \infty$ implique $U \rightarrow B (\leq \infty)$ et réciproquement, alors on a l'égalité

$$\int_a^\infty f(x) dx = \int_b^B f(x(u)) x'(u) du$$

à condition qu'au moins l'une des intégrales écrites existe (alors la deuxième existe aussi).

§ 11.2. Intégrales Impropres de deuxième et de troisième espèce

11.12. Intégrales impropres de deuxième espèce.

a. Soit $f(x)$ une fonction complexe donnée sur un intervalle $[a, b]$ et intégrable sur tout intervalle de la forme $[a + \varepsilon, b]$, mais pas forcément intégrable sur tout l'intervalle $[a, b]$ (elle peut, par

exemple, être non bornée sur $[a, b]$). Nous tenons à attribuer un sens à « l'intégrale impropre de deuxième espèce »

$$\int_a^b f(x) dx. \quad (1)$$

Pour le faire, considérons la fonction de ϵ :

$$I(\epsilon) = \int_{a+\epsilon}^b f(x) dx.$$

Si cette fonction a une limite finie que nous notons I , lorsque $\epsilon \rightarrow 0$, alors on dit que l'intégrale impropre de deuxième espèce (1) *converge*, et on lui attribue la valeur

$$I = \lim_{\epsilon \rightarrow 0} I(\epsilon) = \lim_{\epsilon \rightarrow 0} \int_{a+\epsilon}^b f(x) dx.$$

Si la fonction $I(\epsilon)$ n'a pas de limite pour $\epsilon \rightarrow 0$, on dit que l'intégrale impropre (1) *diverge*, et on ne lui attribue aucune valeur.

b. Si l'intégrale

$$I = \int_a^b f(x) dx \quad (2)$$

existe en tant qu'intégrale définie usuelle, alors elle existe au sens de la nouvelle définition et a la même valeur I . En effet, si l'intégrale (2) existe au sens usuel, alors la fonction $f(x)$ est bornée de même que ses parties réelle et imaginaire (9.15.c); soit, par exemple, $|f(x)| < C$. Pour tout $\epsilon > 0$, on a

$$\int_a^b f(x) dx = \int_a^{a+\epsilon} f(x) dx + \int_{a+\epsilon}^b f(x) dx \quad (3)$$

et

$$\left| \int_a^{a+\epsilon} f(x) dx \right| \leq C\epsilon$$

Il en résulte que la limite de la deuxième intégrale dans le second membre de (3), pour $\epsilon \rightarrow 0$, existe et est égale à l'intégrale dans le premier membre, ce que nous cherchions.

c. La définition a ressemblé à la définition de l'intégrale impropre de première espèce (c.-à-d. à la limite supérieure infinie). En effet, l'intégrale de deuxième espèce se réduit directement à l'intégrale de première espèce par la substitution $x - a = \frac{1}{u}$. Par conséquent,

toute la théorie des intégrales impropres de deuxième espèce peut être déduite de celle des intégrales de première espèce. On peut aussi bien construire ces théories parallèlement. Nous laissons au lecteur le soin de formuler, pour l'intégrale de deuxième espèce, le critère de Cauchy, le critère de comparaison, la définition et le théorème de la convergence absolue, les critères de Leibniz et d'Abel-Dirichlet. Signalons seulement certaines particularités des intégrales de deuxième espèce qui sont essentielles dans les applications.

11.22.a. Pour se servir du critère de comparaison, il faut avoir à sa disposition une vaste collection d'intégrales modèles. En ce qui concerne les intégrales de deuxième espèce, les intégrales les plus usitées de ce type sont

$$\int_a^b \frac{dx}{(x-a)^\lambda}. \quad (1)$$

On a

$$\int_{a+\varepsilon}^b \frac{dx}{(x-a)^\lambda} = \begin{cases} \frac{1}{1-\lambda} \left[\frac{1}{(b-a)^{\lambda-1}} - \frac{1}{\varepsilon^{\lambda-1}} \right] & (\lambda \neq 1), \\ \ln \frac{b-a}{\varepsilon} & (\lambda = 1); \end{cases}$$

l'intégrale de la forme (1) converge donc pour $\lambda < 1$ et diverge pour $\lambda \geq 1$.

b. En appliquant le critère de comparaison on obtient: si, pour $x \in (a, b)$, on a

$$\frac{c_1}{(x-a)^\lambda} \leq f(x) \leq \frac{c_2}{(x-a)^\lambda},$$

alors l'intégrale 11.21. (1) converge pour $\lambda < 1$ et diverge pour $\lambda \geq 1$.

a. Considérons, en particulier, l'intégrale

$$\int_0^b |\ln x|^\gamma dx. \quad (2)$$

Pour x petit, on a $|\ln x| \leq c|x|^{-\alpha}$ quel que soit $\alpha > 0$ (5.57 (3)).

On peut poser $\alpha = \frac{\gamma}{2}$; en comparant ensuite l'intégrale (2) à l'intégrale

$$\int_0^c x^{-\frac{1}{2}} dx$$

qui converge en vertu de a, on voit que l'intégrale (2) converge pour tout γ .

11.23. Intégrales impropres de troisième espèce. Soit $f(x)$ une fonction complexe donnée sur un intervalle (a, b) fini ou non (c.-à-d. que les valeurs $a = -\infty$, $b = +\infty$ sont admissibles). Nous disons que c est un point *singulier* de la fonction $f(x)$ dans les cas suivants:

a) le point $c = -\infty$ ou $c = +\infty$ est singulier s'il représente une extrémité de l'intervalle (a, b) ;

b) un point c à l'intérieur de l'intervalle (a, b) est singulier si la fonction $f(x)$ n'est intégrable au sens ordinaire dans aucun voisinage de ce point; l'extrémité c est un point singulier si $f(x)$ n'est intégrable dans aucune intersection de l'intervalle (a, b) avec un voisinage de c .

Supposons que la fonction $f(x)$ possède au plus un nombre fini de points singuliers et que, à l'extérieur de ces points, elle soit continue ou continue par morceaux. Nous tenons à attribuer un sens à « l'intégrale impropre de troisième espèce »

$$\int_a^b f(x) dx. \quad (1)$$

En fixant, entre deux points singuliers successifs quelconques c_i , c_{i+1} , un point non singulier p_i , on aboutit à un ensemble des intervalles de la forme (c_i, p_i) et (p_i, c_{i+1}) ayant chacun un seul point singulier à l'extrémité. Définissons d'abord les intégrales de la fonction $f(x)$ sur ces intervalles comme intégrales impropres de 1-ière ou 2-ième espèce:

$$\int_{c_i}^{p_i} f(x) dx = \lim_{x \rightarrow c_i} \int_x^{p_i} f(x) dx, \quad \int_{p_i}^{c_{i+1}} f(x) dx = \lim_{x \rightarrow c_{i+1}} \int_{p_i}^x f(x) dx.$$

On n'attribue un sens à l'intégrale (1) que si toutes les intégrales impropres ainsi obtenues convergent; notamment, on pose

$$\int_a^b f(x) dx = \sum \left\{ \int_{c_i}^{p_i} f(x) dx + \int_{p_i}^{c_{i+1}} f(x) dx \right\}. \quad (2)$$

Il faut prouver que le résultat ne dépend pas du choix de points p_i . Faisons-le pour l'un des intervalles $[c_i, c_{i+1}]$. Soit $c_i < p_i < q_i < c_{i+1}$. Alors on a

$$\begin{aligned} \int_{c_i}^{p_i} f(x) dx + \int_{p_i}^{c_{i+1}} f(x) dx &= \lim_{x \rightarrow c_i} \int_x^{p_i} f(x) dx + \lim_{y \rightarrow c_{i+1}} \int_{p_i}^y f(x) dx = \\ &= \lim_{x \rightarrow c_i} \int_x^{p_i} f(x) dx + \lim_{y \rightarrow c_{i+1}} \left\{ \int_{p_i}^{q_i} f(x) dx + \int_{q_i}^y f(x) dx \right\} = \lim_{x \rightarrow c_i} \int_x^{p_i} f(x) dx + \int_{p_i}^{q_i} f(x) dx + \lim_{y \rightarrow c_{i+1}} \int_{q_i}^y f(x) dx = \end{aligned}$$

$$\begin{aligned}
&= \lim_{X \rightarrow c_i} \left\{ \int_X^{p_i} + \int_{p_i}^{q_i} \right\} + \lim_{Y \rightarrow c_{i+1}} \int_{q_i}^Y = \lim_{X \rightarrow c_i} \int_X^{q_i} + \lim_{Y \rightarrow c_{i+1}} \int_{q_i}^Y = \\
&= \int_{c_i}^{q_i} f(x) dx + \int_{q_i}^{c_{i+1}} f(x) dx,
\end{aligned}$$

ce qu'il fallait démontrer. (Lors des calculs intermédiaires, on n'explicitait pas les expressions à intégrer.)

11.24. Exemple. Considérons l'intégrale

$$\int_{-\infty}^{\infty} \frac{dx}{|x-c_1|^{\alpha_1} \dots |x-c_n|^{\alpha_n}}, \quad -\infty < c_1 < \dots < c_n < \infty. \quad (1)$$

Les points singuliers sont $-\infty, c_1, \dots, c_n, \infty$.

Les intégrales $\int_{c_i}^{p_i}$ et $\int_{p_{i-1}}^{c_i}$ convergent si, et seulement si, $\alpha_i < 1$

(les autres facteurs étant bornés au voisinage du point c_i). Les

intégrales $\int_{-\infty}^{p_0}$ et $\int_{p_n}^{\infty}$ convergent si, et seulement si, $\alpha_1 + \dots + \alpha_n >$

> 1 . Donc, l'intégrale (1) n'existe que pour

$$\alpha_1 < 1, \dots, \alpha_n < 1, \alpha_1 + \dots + \alpha_n > 1.$$

Notons que l'intégrale

$$\int_{-\infty}^{\infty} \frac{dx}{|x-c|^{\alpha}}$$

qui s'obtient de (1), lorsque $n=1$, ne converge pas quelle que soit la valeur de α . Il en est de même des intégrales

$$\int_{-\infty}^c \frac{dx}{(c-x)^{\alpha}}, \quad \int_c^{\infty} \frac{dx}{(x-c)^{\alpha}}.$$

§ 11.3. Calcul d'Intégrales impropres à l'aide de résidus

11.31. Intégrale d'une fonction rationnelle. Considérons l'intégrale impropre d'une fonction rationnelle, c.-à-d. du rapport de deux polynômes $P(x)$ et $Q(x)$ (à coefficients com-

plexes) :

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx. \quad (1)$$

Elle converge si le dénominateur n'a pas de racines réelles et le degré du dénominateur dépasse au moins de deux unités celui du numérateur (11.13.c).

Comment trouver la valeur de cette intégrale ?

On peut certainement calculer l'intégrale indéfinie de la fonction rationnelle $\frac{P(x)}{Q(x)}$ selon les règles 9.42 et substituer ensuite les limites $-\infty$ et ∞ . Mais il s'avère qu'il vaut mieux, parfois, d'appliquer les méthodes liées à la nature analytique de la fonction $\frac{P(z)}{Q(z)}$.

La fonction $\frac{P(z)}{Q(z)}$ de la variable complexe z est analytique partout dans le plan des z , à l'exception d'un nombre fini de points qui sont les racines du dénominateur. Considérons dans le demi-plan supérieur le contour fermé lisse par morceaux L formé du segment $[-R, R]$ de l'axe réel et de la demi-circonférence

$$C_R^+ = \{z = Re^{i\varphi}, 0 \leq \varphi \leq \pi\},$$

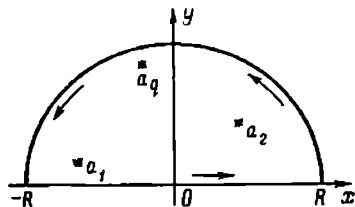


Fig. 11.5.

R étant si grand qu'il n'y a pas de racines du dénominateur dans le demi-plan supérieur à l'extérieur du demi-cercle construit.

A son intérieur, il y a un certain nombre de racines du dénominateur, par exemple $a_1, \dots, a_q$ (fig. 11.5). En vertu de la formule 10.43 (1), on a

$$\int_{-R}^R \frac{P(x)}{Q(x)} dx + \int_{C_R^+} \frac{P(z)}{Q(z)} dz = 2\pi i \sum_{j=1}^q \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=a_j}. \quad (2)$$

Faisons tendre R vers ∞ . Vu la condition imposée aux degrés des polynômes $P(z)$ et $Q(z)$, on a $\left| \frac{P(z)}{Q(z)} \right| \leq \frac{A}{R^2}$ sur la demi-circonférence C_R^+ , A étant une constante ; donc

$$\left| \int_{C_R^+} \frac{P(z)}{Q(z)} dz \right| \leq \frac{A}{R^2} \pi R = \frac{\pi A}{R}.$$

Il en résulte que l'intégrale

$$\int_{-R}^R \frac{P(x)}{Q(x)} dx \quad (3)$$

a la quantité (2) pour limite lorsque $R \rightarrow \infty$. L'intégrale (1) étant convergente, elle doit se confondre avec la limite de l'intégrale (3). Définitivement, on a

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx = 2\pi i \sum_{j=1}^q \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=a_j} \quad (4)$$

a. Si les racines $a_1, \dots, a_q$ sont simples, alors, d'après la formule 10.42 (1), on a

$$\operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=a_j} = \frac{P'(a_j)}{Q'(a_j)} ;$$

par conséquent,

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx = 2\pi i \sum_{j=1}^q \frac{P(a_j)}{Q'(a_j)} \quad (5)$$

Exemple :

$$\int_{-\infty}^{\infty} \frac{dx}{x^2+1} = 2\pi i \frac{1}{2z} \Big|_{z=i} = \pi.$$

b. Si les racines $a_1, \dots, a_q$ sont multiples, de sorte que la racine a_j est de multiplicité k_j , alors on doit appliquer la formule plus générale 10.42 (2) à la place de 10.42 (1) :

$$\operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=a_j} = \frac{1}{(k_j-1)!} \left[(z-a_j)^{k_j} \frac{P(z)}{Q(z)} \right]_{z=a_j}^{(k_j-1)}.$$

de sorte que

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx = 2\pi i \sum_{j=1}^q \frac{1}{(k_j-1)!} \left[(z-a_j)^{k_j} \frac{P(z)}{Q(z)} \right]_{z=a_j}^{(k_j-1)}.$$

Exemple :

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{(x^2+1)^2} &= 2\pi i \left[(z-i)^2 \frac{1}{(z^2+1)^2} \right]_{z=i}' = \\ &= 2\pi i \left[\frac{1}{(z+i)^2} \right]_{z=i}' = \frac{-4\pi i}{(z+i)^3} \Big|_{z=i} = -4\pi i \frac{1}{8i^3} = \frac{\pi}{2}. \end{aligned}$$

c. Remarque. Nous avons réduit l'intégrale $\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx$ à la

somme des résidus (multipliée par $2\pi i$) de la fonction $\frac{P(z)}{Q(z)}$ dans le demi-plan supérieur en considérant le contour L formé du segment $[-R, R]$ et de la demi-circonférence $C_R^+ = \{z = Re^{i\varphi}, 0 \leq \varphi \leq \pi\}$. On peut faire de même avec la contour L^- composé par le segment $[R, -R]$ (parcours de droite à gauche) et par la demi-circonférence $C_R^- = \{z = Re^{i\varphi}, \pi \leq \varphi \leq 2\pi\}$ dans le demi-plan inférieur; on a

$$2\pi i \sum_{h=1}^r \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=b_h} = \oint_{L^-} \frac{P(z)}{Q(z)} dz = \int_R^{-R} \frac{P(x)}{Q(x)} dx + \oint_{C_R^-} \frac{P(z)}{Q(z)} dz,$$

$b_1, \dots, b_r$ étant les racines du polynôme $Q(z)$ appartenant au demi-plan inférieur.

En passant à la limite pour $R \rightarrow \infty$ on trouve

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx = \lim_{R \rightarrow \infty} \int_R^{-R} \frac{P(x)}{Q(x)} dx = -2\pi i \sum_{h=1}^r \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=b_h}. \quad (7)$$

Les résultats obtenus ici et dans (4) ont les formes différentes. En réalité, ils se confondent parce que la somme des résidus (multipliée par $2\pi i$) de la fonction $\frac{P(z)}{Q(z)}$ pour toutes les racines de $Q(z)$ est nulle.

On peut le voir directement. On sait que la somme mentionnée des résidus coïncide avec l'intégrale

$$2\pi i \oint_{|z|=R} \frac{P(z)}{Q(z)} dz \quad (8)$$

prise le long de la circonférence d'un rayon R si grand qu'elle contienne toutes les racines de $Q(z)$ à son intérieur. Cette intégrale ne dépend pas de R et admet l'estimation

$$\left| \oint_{|z|=R} \frac{P(z)}{Q(z)} dz \right| \leq \max_{|z|=R} \left| \frac{P(z)}{Q(z)} \right| \cdot 2\pi R \leq \frac{A}{R^2} \cdot 2\pi R \rightarrow 0.$$

L'intégrale (8) est donc nulle, d'où

$$\sum_{j=1}^q \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=a_j} + \sum_{h=1}^r \operatorname{Res} \left[\frac{P(z)}{Q(z)} \right]_{z=b_h} = 0.$$

11.32.a. Intégrales de Fourier. Considérons les intégrales usitées de la forme

$$\int_{-\infty}^{\infty} f(x) e^{i\sigma x} dx, \quad (1)$$

$$\int_{-\infty}^{\infty} f(x) \cos \sigma x dx, \quad (2)$$

$$\int_{-\infty}^{\infty} f(x) \sin \sigma x dx \quad (3)$$

qui contiennent un paramètre réel σ ; elles s'appellent *intégrales de Fourier*.

Si la condition

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty$$

est satisfaite, alors toutes les trois intégrales de Fourier convergent absolument. Si, pour $|x| \rightarrow \infty$, la fonction $f(x)$ est réelle et tend vers zéro d'une façon monotone, alors les intégrales (2) et (3) convergent pour $\sigma \neq 0$ (cf. 11.16.a), mais, dans le cas général, non absolument. Si, de plus, $f(-x) \equiv f(x)$ (la fonction $f(x)$ est paire), alors l'intégrale (3) est nulle, et si $f(-x) = -f(x)$ ($f(x)$ est impaire), c'est l'intégrale (2) qui est nulle. En outre, on a la relation évidente

$$\int_{-\infty}^{\infty} f(x) e^{i\sigma x} dx = \int_{-\infty}^{\infty} f(x) \cos \sigma x dx + i \int_{-\infty}^{\infty} f(x) \sin \sigma x dx.$$

de sorte que, dans le cas où $f(x)$ est réelle, les intégrales (2) et (3) représentent les parties réelle et imaginaire respectivement de l'intégrale (1).

b. La théorie générale des intégrales de Fourier sera exposée dans le chapitre 15 (troisième partie). Ici nous nous bornons au calcul des intégrales de Fourier pour certaines classes de fonctions. Les méthodes d'intégration le long d'un contour sont en l'occurrence très utiles.

Soit $f(x) = \frac{P(x)}{Q(x)}$ une fonction rationnelle et supposons que le degré du polynôme $Q(x)$ dépasse au moins d'une unité celui de $P(x)$ et que le dénominateur ne s'annule pour aucun x réel. Dans ce cas, les intégrales (1)-(3) convergent (11.17.b); on va les calculer.

Considérons les racines $a_1, \dots, a_m$ du polynôme $Q(x)$ dans le

demi-plan supérieur. Construisons le contour fermé L^+ du segment $[-R, R]$ de l'axe réel et de la demi-circonférence $C_R^+ = \{z = Re^{i\varphi}, 0 \leq \varphi \leq \pi\}$. Alors, d'après 10.43 (1), on a

$$\oint_{L^+} \frac{P(z)}{Q(z)} e^{i\sigma z} dz = \int_{-R}^R \frac{P(x)}{Q(x)} e^{i\sigma x} dx + \int_{C_R^+} \frac{P(z)}{Q(z)} e^{i\sigma z} dz = \\ = 2\pi i \sum_{j=1}^m \operatorname{Res} \left[\frac{P(z)}{Q(z)} e^{i\sigma z} \right]_{z=a_j}. \quad (4)$$

Montrons que, pour $\sigma > 0$,

$$\lim_{R \rightarrow \infty} \int_{C_R^+} \frac{P(z)}{Q(z)} e^{i\sigma z} dz = 0. \quad (5)$$

Si $y \geq 0$, $\sigma \geq 0$, alors $|e^{i\sigma z}| = |e^{i\sigma x - \sigma y}| = e^{-\sigma y} \leq 1$. Par conséquent, si le degré du polynôme $Q(z)$ dépasse au moins de deux unités celui de $P(z)$, la démonstration de la relation (5) se fait exactement de même que dans 11.31.

c. Si le degré du polynôme $Q(z)$ ne dépasse que d'une unité celui de $P(z)$, le raisonnement 11.31 ne convient pas. Pour ce cas, nous établissons le lemme suivant.

Lemme. Si $\sigma > 0$, on a

$$\int_0^\pi e^{-R\sigma \sin \varphi} d\varphi \leq \frac{c}{R\sigma} \quad (c \text{ constant}).$$

Démonstration. Comme $\sin(\pi - \varphi) = \sin \varphi$, il suffit de considérer l'intégrale

$$\int_0^{\pi/2} e^{-R\sigma \sin \varphi} d\varphi$$

qui fait la moitié de la précédente. Pour $R \geq 1$, on a

$$\int_0^{\pi/2} e^{-R\sigma \sin \varphi} d\varphi = \int_0^{\pi/6} e^{-R\sigma \sin \varphi} d\varphi + \int_{\pi/6}^{\pi/2} e^{-R\sigma \sin \varphi} d\varphi < \\ < \int_0^{\pi/6} \frac{\cos \varphi}{\cos \frac{\pi}{6}} e^{-R\sigma \sin \varphi} d\varphi + \int_{\pi/6}^{\pi/2} e^{-R\sigma \sin \varphi} d\varphi < \\ < \frac{e^{-R\sigma \sin \varphi}}{R\sigma \cos \frac{\pi}{6}} \Big|_{\varphi=\frac{\pi}{6}}^0 + \frac{\pi}{3} e^{-\frac{R\sigma}{2}} < \frac{1}{R\sigma \cos \frac{\pi}{6}} + \frac{\pi}{3} e^{-\frac{R\sigma}{2}} = \\ = \frac{1}{R\sigma} \left(\frac{1}{\cos \frac{\pi}{6}} + \frac{\pi}{3} R\sigma e^{-\frac{R\sigma}{2}} \right) \leq \frac{c}{R\sigma},$$

la fonction ue^{-u} étant bornée pour $u > 0$. Le lemme est démontré.

d. Conséquence (lemme de Jordan). Si une fonction $f(z)$ est définie et analytique dans le demi-plan $\text{Im } z \geq 0$ pour $|z| \geq p$, et si, de plus,

$$\lim_{R \rightarrow \infty} \sup_{|z|=R} |f(z)| = 0,$$

alors, pour $\sigma > 0$, on a

$$\lim_{R \rightarrow \infty} \oint_{C_R^+} f(z) e^{i\sigma z} dz = 0.$$

Démonstration. En appliquant le lemme c on trouve

$$\begin{aligned} \left| \oint_{C_R^+} f(z) e^{i\sigma z} dz \right| &= \left| \int_0^\pi f(z) e^{i\sigma z} e^{-\sigma R \sin \varphi} R e^{i\varphi} d\varphi \right| \leq \\ &\leq \sup_{z \in C_R^+} |f(z)| \cdot R \cdot \int_0^\pi e^{-R\sigma \sin \varphi} d\varphi \leq \frac{c}{\sigma} \sup_{z \in C_R^+} |f(z)| \rightarrow 0 \text{ pour } R \rightarrow \infty, \end{aligned}$$

ce qu'il fallait démontrer.

e. On continue le calcul de l'intégrale (1) pour $f(z) = \frac{P(z)}{Q(z)}$. Comme cette fonction, pour les $|z|$ assez grands, satisfait à l'inégalité

$$|f(z)| = \left| \frac{P(z)}{Q(z)} \right| \leq \frac{c}{|z|},$$

l'égalité (5) résulte du lemme de Jordan. En passant, dans (4), à la limite lorsque $R \rightarrow \infty$ on obtient l'intégrale de Fourier de la fonction $f(z)$ avec $\sigma > 0$:

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} e^{i\sigma x} dx = 2\pi i \sum_{j=1}^m \text{Res} \left[\frac{P(z)}{Q(z)} e^{i\sigma z} \right]_{z=a_j}.$$

f. Si $\sigma < 0$, le raisonnement ci-dessus n'est plus valable, parce que la fonction $|e^{i\sigma z}|$ croît indéfiniment avec R sur l'arc C_R^+ . Dans ce cas, nous faisons la construction analogue non pas dans le demi-plan supérieur, mais dans le demi-plan inférieur. Désignons par C_R^- la demi-circonférence $\{z = Re^{i\varphi}, \pi \leq \varphi \leq 2\pi\}$ d'un rayon si grand que toutes les racines $b_1, \dots, b_r$ du polynôme $Q(z)$ dans le demi-plan inférieur se trouvent à son intérieur; on a alors:

$$\begin{aligned} \oint_{C_R^-} \frac{P(z)}{Q(z)} e^{i\sigma z} dz &= \int_{-R}^{-R} \frac{P(x)}{Q(x)} e^{i\sigma x} dx + \oint_{C_R^-} \frac{P(z)}{Q(z)} e^{i\sigma z} dz = \\ &= 2\pi i \sum_{h=1}^r \text{Res} \left[\frac{P(z)}{Q(z)} e^{i\sigma z} \right]_{z=b_h}. \end{aligned}$$

L'intégrale prise le long de C_R^- admet l'estimation

$$\left| \int_{C_R^-} \frac{P(z)}{Q(z)} e^{i\sigma z} dz \right| = \left| \int_{\pi}^{2\pi} \frac{P(z)}{Q(z)} e^{i\sigma z} e^{-\sigma v} R e^{i\varphi} d\varphi \right| \leq \\ \leq \max_{z \in C_R^-} \left| \frac{P(z)}{Q(z)} \right| \cdot R \cdot \int_0^{\pi} e^{R\sigma \sin \varphi} d\varphi$$

et, en vertu du lemme, tend vers zéro pour $R \rightarrow \infty$, d'où

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} e^{i\sigma x} dx = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{P(x)}{Q(x)} e^{i\sigma x} dx = \\ = -2\pi i \sum \operatorname{Res} \left[\frac{P(z)}{Q(z)} e^{i\sigma z} \right]_{z=b_k},$$

ce qui représente la valeur de l'intégrale de Fourier de la fonction $\frac{P(x)}{Q(x)}$ pour $\sigma < 0$.

g. Pour $\sigma = 0$, l'intégrale de Fourier (1) devient intégrale de la fonction rationnelle $\frac{P(x)}{Q(x)}$, et si la différence des degrés du dénominateur et du numérateur fait un, elle n'existe pas.

h. Exemple (intégrale de Laplace):

$$\int_{-\infty}^{\infty} \frac{x e^{i\sigma x}}{x^2 + 1} dx = \begin{cases} 2\pi i \operatorname{Res} \left[\frac{z e^{i\sigma z}}{z^2 + 1} \right]_{z=i} = i\pi e^{-\sigma} & (\sigma > 0), \\ -2\pi i \operatorname{Res} \left[\frac{z e^{i\sigma z}}{z^2 + 1} \right]_{z=-i} = -i\pi e^{\sigma} & (\sigma < 0). \end{cases}$$

Notons que l'intégrale indéfinie de la fonction $\frac{x e^{i\sigma x}}{x^2 + 1}$ ne s'exprime pas en fonctions élémentaires.

11.33. Intégrales spéciales de Fourier. Il arrive que l'une ou l'autre des intégrales de Fourier 11.32 (2) et (3) existe bien que la fonction $f(x)$ ait des singularités sur l'axe même des x , si ces singularités sont compensées par les zéros correspondants des fonctions $\sin \sigma x$ ou $\cos \sigma x$. Considérons, par exemple, l'intégrale usitée en physique mathématique

$$\int_{-\infty}^{\infty} \frac{\sin \sigma x}{x} dx$$

(sa convergence résulte du critère de Leibniz 11.16). Dans le présent cas, l'intégrale de Fourier $\int_{-\infty}^{\infty} \frac{e^{i\sigma x}}{x} dx$ n'existe pas.

Considérons le contour L formé successivement par le segment $[-R, -\varepsilon]$ de l'axe réel, la petite demi-circonférence

$$C_\varepsilon = \{z = \varepsilon e^{i\varphi}, 0 \leq \varphi \leq \pi\},$$

le segment $[\varepsilon, R]$ de l'axe réel et la grande demi-circonférence

$$\Gamma_R = \{z = R e^{i\varphi}, 0 \leq \varphi \leq \pi\}$$

(fig. 11.6). La fonction $e^{i\sigma z}/z$ n'ayant aucune singularité à l'intérieur

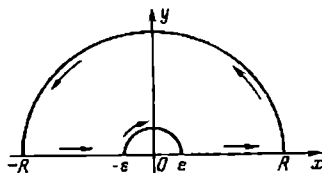


Fig. 11.6.

du contour L ni sur le contour même, on a

$$\oint_L \frac{e^{i\sigma z}}{z} dz = \int_{-R}^{-\varepsilon} \frac{e^{i\sigma x}}{x} dx + \int_{C_\varepsilon} \frac{e^{i\sigma z}}{z} dz + \int_{\varepsilon}^R \frac{e^{i\sigma x}}{x} dx + \int_{\Gamma_R} \frac{e^{i\sigma z}}{z} dz = 0.$$

En appliquant le lemme de Jordan on a, pour $\sigma > 0$:

$$\lim_{R \rightarrow \infty} \int_{\Gamma_R} \frac{e^{i\sigma z}}{z} dz = 0.$$

Ensuite,

$$\begin{aligned} \int_{-R}^{-\varepsilon} \frac{e^{i\sigma x}}{x} dx + \int_{\varepsilon}^R \frac{e^{i\sigma x}}{x} dx &= \left[\int_{-R}^{-\varepsilon} \frac{\cos \sigma x}{x} dx + \int_{\varepsilon}^R \frac{\cos \sigma x}{x} dx \right] + \\ &+ i \left[\int_{-R}^{-\varepsilon} \frac{\sin \sigma x}{x} dx + \int_{\varepsilon}^R \frac{\sin \sigma x}{x} dx \right] = i \left[\int_{-R}^{-\varepsilon} \frac{\sin \sigma x}{x} dx + \int_{\varepsilon}^R \frac{\sin \sigma x}{x} dx \right], \end{aligned}$$

car les intégrales de la fonction $\frac{\cos \sigma x}{x}$ sur les intervalles $[-R, -\varepsilon]$ et $[\varepsilon, R]$ se réduisent en vertu de l'imparité de la fonction. Enfin,

$$\int_{C_\varepsilon} \frac{e^{i\sigma z}}{z} dz = \int_{C_\varepsilon} \frac{dz}{z} + \int_{C_\varepsilon} \frac{e^{i\sigma z} - 1}{z} dz = \int_{\varphi=\pi}^0 \frac{te^{i\sigma \varepsilon e^{i\varphi}} d\varphi}{\varepsilon e^{i\varphi}} + R_\varepsilon = -i\pi + R_\varepsilon,$$

où

$$|R_\varepsilon| = \left| \int_{C_\varepsilon} \frac{e^{i\sigma z} - 1}{z} dz \right| \leq \max_{z \in C_\varepsilon} \left| \frac{e^{i\sigma z} - 1}{z} \right| \varepsilon \pi \leq M\varepsilon,$$

puisque la fonction $\frac{e^{i\sigma z}-1}{z}$ est continue, donc bornée à proximité de $z=0$. Lorsque z tend vers 0 et R vers ∞ , on obtient

$$\int_{-\infty}^{\infty} \frac{\sin \sigma x}{x} dx = \pi. \quad (1)$$

Si $\sigma < 0$, alors $\sin \sigma x = -\sin(-\sigma)x$, où $-\sigma > 0$; par conséquent

$$\int_{-\infty}^{\infty} \frac{\sin \sigma x}{x} dx = - \int_{-\infty}^{\infty} \frac{\sin(-\sigma)x}{x} dx = -\pi.$$

Il en résulte que (Euler, 1781)

$$\int_0^{\infty} \frac{\sin \sigma x}{x} dx = \begin{cases} \frac{\pi}{2} & \text{pour } \sigma > 0, \\ -\frac{\pi}{2} & \text{pour } \sigma < 0. \end{cases} \quad (2)$$

Une méthode analogue permet de calculer les intégrales plus compliquées

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} \frac{\sin x}{x} dx,$$

où $Q(x) \neq 0$ (cf. exercice 5).

§ 11.4. Intégrales Impropres fonctions d'un paramètre

11.41. L'intégrale de Fourier représente un exemple d'intégrale impropre fonction d'un paramètre. Considérons une intégrale impropre arbitraire, de première espèce par exemple, qui dépend d'un paramètre réel σ :

$$I(\sigma) = \int_a^{\infty} y(t, \sigma) dt. \quad (1)$$

Nous supposons qu'elle converge pour toute valeur de σ d'un intervalle $\alpha \leq \sigma \leq \beta$; nous nous intéressons aux propriétés du résultat de l'intégration $I(\sigma)$ en tant que fonction de σ . Notamment, nous tenons à trouver les réponses aux questions suivantes:

(a) Si la fonction $y(t, \sigma)$ est continue par rapport à σ , la fonction $I(\sigma)$ est-elle continue?

(b) Est-ce que la formule d'intégration par rapport au paramètre

$$\int_{\alpha}^{\beta} \left\{ \int_a^{\infty} y(t, \sigma) dt \right\} d\sigma = \int_a^{\infty} \left\{ \int_{\alpha}^{\beta} y(t, \sigma) d\sigma \right\} dt$$

a lieu?

(c) La fonction $y(t, \sigma)$ étant dérivable par rapport à σ , est-ce que la formule de dérivation

$$\frac{d}{d\sigma} \int_a^\infty y(t, \sigma) dt = \int_a^\infty \frac{d}{d\sigma} y(t, \sigma) dt$$

est valable?

En tant qu'exemple, considérons l'intégrale

$$I(\sigma) = \int_0^\infty \frac{\sin t\sigma}{t} dt$$

convergeant pour tout σ et calculée dans 11.33; on a vu que

$$I(\sigma) = \begin{cases} \frac{\pi}{2} & \text{pour } \sigma > 0, \\ -\frac{\pi}{2} & \text{pour } \sigma < 0. \end{cases}$$

De plus, il est évident que $I(0) = 0$.

On voit que bien que $\frac{\sin t\sigma}{t}$ soit continue par rapport au paramètre σ , le résultat de l'intégration $I(\sigma)$ a une discontinuité pour $\sigma = 0$. Par conséquent, la réponse à la question (a) est, en général, négative. Ensuite, dans le présent exemple nous avons $\frac{d}{d\sigma} \frac{\sin t\sigma}{t} = \cos t\sigma$, et la formule de dérivation (c) n'a pas lieu car l'intégrale dans le second membre est divergente. Bien que la formule d'intégration (b) soit valable pour notre exemple, nous comprenons que cela est dû aux propriétés particulières de la fonction $\frac{\sin t\sigma}{t}$. Un exemple de la réponse négative à la question (b) est donné dans l'exercice 6. Pour que la réponse aux questions (a)-(c) soit positive, il faut imposer des conditions supplémentaires au caractère de convergence de l'intégrale impropre (1).

Nous ne ferons ici les constructions nécessaires que pour les intégrales de première espèce, mais tous les définitions et théorèmes peuvent être étendus aux intégrales de deuxième et troisième espèce, ce que le lecteur vérifiera aisément.

11.42. Définition. Une intégrale impropre fonction d'un paramètre

$$I(\sigma) = \int_a^\infty y(t, \sigma) dt \quad (1)$$

qui converge pour toute valeur de σ d'un ensemble S est dite *uniformément convergente sur S* (ou *par rapport à $\sigma \in S$*) si, quel que

soit $\varepsilon > 0$, il existe un N tel que, pour tout $p \geq N$ et pour tout $\sigma \in S$, on a

$$\left| \int_p^\infty y(t, \sigma) dt \right| < \varepsilon.$$

Choisissons arbitrairement les nombres $h_n \nearrow \infty$ et écrivons la suite des fonctions

$$I_n(\sigma) = \int_a^{a+h_n} y(t, \sigma) dt \quad (n = 1, 2, \dots). \quad (2)$$

Si l'intégrale (1) converge uniformément sur l'ensemble S , alors la suite (2) converge uniformément sur le même ensemble vers sa limite qui est l'intégrale (1).

11.43. Théorème. *Si S est un espace métrique, la fonction $y(t, \sigma)$ est uniformément continue sur tout produit $S \times [a, a+h]$, $0 < h < \infty$, et l'intégrale 11.42 (1) converge uniformément sur S , alors la fonction $I(\sigma)$ est continue sur l'ensemble S .*

Démonstration. D'après le théorème 9.81, la fonction $I_n(\sigma)$ est continue pour $\sigma \in S$. D'après le théorème 5.96, la fonction $I(\sigma)$ est aussi continue pour $\sigma \in S$ comme limite d'une suite uniformément convergente de fonctions continues. Le théorème est démontré.

11.44. Théorème. *Si une fonction $y(t, \sigma)$ est continue par rapport à l'ensemble des variables (t, σ) sur tout rectangle*

$$S = \{(t, \sigma) : a \leq t \leq a+h, 0 < h < \infty; \alpha \leq \sigma \leq \beta\},$$

et que l'intégrale 11.42 (1) converge pour $\alpha \leq \sigma \leq \beta$ uniformément par rapport à σ , alors la fonction

$$Y(t) = \int_\alpha^\beta y(t, \sigma) d\sigma$$

possède son intégrale impropre convergente sur l'intervalle $a \leq t < \infty$ et l'on a

$$\int_a^\infty \left\{ \int_\alpha^\beta y(t, \sigma) dt \right\} d\sigma = \int_\alpha^\beta \left\{ \int_a^\infty y(t, \sigma) dt \right\} d\sigma. \quad (1)$$

Démonstration. D'après le théorème 9.72, la convergence uniforme des fonctions continues (11.43) $I_n(\sigma)$ vers $I(\sigma)$ implique

$$\int_\alpha^\beta I(\sigma) d\sigma = \lim_{n \rightarrow \infty} \int_\alpha^\beta I_n(\sigma) d\sigma. \quad (2)$$

Or, d'après le théorème 9.82, on a

$$\int_{\alpha}^{\beta} I_n(\sigma) d\sigma = \int_{\alpha}^{\beta} \left\{ \int_a^{a+h_n} y(t, \sigma) dt \right\} d\sigma = \int_a^{a+h_n} \left\{ \int_{\alpha}^{\beta} y(t, \sigma) d\sigma \right\} dt.$$

L'égalité (2) prend donc la forme

$$\int_{\alpha}^{\beta} I(\sigma) d\sigma = \lim_{n \rightarrow \infty} \int_a^{a+h_n} \left\{ \int_{\alpha}^{\beta} y(t, \sigma) d\sigma \right\} dt. \quad (3)$$

L'existence de la limite du second membre pour $h_n \nearrow \infty$ signifie l'existence de l'intégrale impropre $\int_a^{\infty} \left\{ \int_{\alpha}^{\beta} y(t, \sigma) d\sigma \right\} dt$ (11.12. b).

Par conséquent, (1) résulte de (3). Le théorème est démontré.

11.45.a. Théorème. Si la dérivée partielle $y'_{\sigma}(t, \sigma)$ est continue par rapport à l'ensemble (t, σ) dans tout rectangle $\alpha \leq \sigma \leq \beta$, $a \leq t \leq a+h$, $0 < h < \infty$, l'intégrale $\int_a^{\infty} y'_{\sigma}(t, \sigma) dt$ converge pour $\alpha \leq \sigma \leq \beta$ uniformément par rapport à σ et l'intégrale $\int_a^{\infty} y(t, \sigma) dt$ converge, alors la fonction

$$I(\sigma) = \int_a^{\infty} y(t, \sigma) dt$$

existe pour $\alpha \leq \sigma \leq \beta$, est dérivable par rapport à σ , et l'on a

$$\frac{d}{d\sigma} \int_a^{\infty} y(t, \sigma) dt = \int_a^{\infty} y'_{\sigma}(t, \sigma) dt. \quad (1)$$

Démonstration. En vertu du théorème 9.84,

$$I'_n(\sigma) = \frac{d}{d\sigma} \int_a^{a+h_n} y(t, \sigma) dt = \int_a^{a+h_n} y'_{\sigma}(t, \sigma) dt.$$

Par hypothèse, la suite des fonctions $I'_n(\sigma)$ converge uniformément sur l'intervalle $\alpha \leq \sigma \leq \beta$, et la suite des fonctions $I_n(\sigma)$ converge au point $\sigma = \alpha$. D'après le théorème 9.77, la suite $I_n(\sigma)$ converge uniformément sur l'intervalle $\alpha \leq \sigma \leq \beta$, sa limite $I(\sigma)$ est une fonction dérivable pour $\sigma \in [\alpha, \beta]$, et l'on a

$$I'(\sigma) = \lim_{n \rightarrow \infty} I'_n(\sigma).$$

On a donc démontré l'existence et la dérivabilité par rapport à σ de l'intégrale

$$I(\sigma) = \int_a^\infty y(t, \sigma) dt = \lim_{n \rightarrow \infty} I_n(\sigma)$$

ainsi que l'égalité (1).

b. Pour les fonctions analytiques, l'énoncé du théorème a se modifie.

Supposons que la fonction $y(t, s)$ soit définie pour $t \in (a, \infty)$ et $s \in G$, où G est un domaine dans le plan $s = \sigma + i\tau$; supposons ensuite que $y(t, s)$ soit analytique par rapport à $s \in G$, pour tout t , et continue par rapport à l'ensemble (t, s) dans tout « cylindre » de la forme $a \leq t \leq a + h$, $\{s: |s - s_0| \leq \rho\} \subset G$ ($0 < h < \infty$). Nous dirons que l'intégrale

$$\int_a^\infty y(t, s) dt, \quad s \in G, \quad (2)$$

converge uniformément à l'intérieur de G si elle converge pour tout $s \in G$ et si, pour tout compact $Q \subset G$ et pour tout $\varepsilon > 0$, il existe un $N = N(\varepsilon, Q)$ tel que

$$\left| \int_N^\infty y(t, s) dt \right| < \varepsilon$$

quel que soit $s \in Q$.

Théorème. Si l'intégrale (2) converge uniformément à l'intérieur de G , alors la fonction

$$I(s) = \int_a^\infty y(t, s) dt$$

est analytique dans G et

$$I'(s) = \int_a^\infty y'_s(t, s) dt.$$

Démonstration. D'après le théorème 10.27, la fonction

$$I_n(s) = \int_a^{h_n} y(t, s) dt$$

est analytique par rapport à s , et de plus $I'_n(s) = \int_a^{h_n} y'_s(t, s) dt$.

Ensuite, la fonction $I_n(s)$ converge uniformément à l'intérieur de G vers la fonction $I(s)$, lorsque $n \rightarrow \infty$; donc, en vertu du théorème 10.36, la fonction $I_n(s)$ est analytique dans G . D'après le même théorème 10.36, les fonctions $I_n(s)$ convergent vers la fonction $I'(s)$ dans le domaine G , d'où

$$I'(s) = \lim_{n \rightarrow \infty} \int_a^{h_n} y'_s(t, s) dt = \int_a^\infty y'_s(t, s) dt,$$

ce qu'il fallait démontrer.

11.46. Critère de Cauchy de convergence uniforme des intégrales. Une intégrale impropre

$$\int_a^\infty y(t, \sigma) dt \quad (1)$$

avec le paramètre σ parcourant un ensemble S converge uniformément sur S si, et seulement si, pour tout $\varepsilon > 0$, il existe un nombre N tel que, quels que soient $p \geq N$, $q \geq N$, $\sigma \in S$, on a

$$\left| \int_p^q y(t, \sigma) dt \right| \leq \varepsilon. \quad (2)$$

Démonstration. Supposons que l'intégrale (1) converge uniformément sur S . Pour un $\varepsilon > 0$ donné, trouvons un nombre N de façon à avoir

$$\left| \int_p^\infty y(t, \sigma) dt \right| < \frac{\varepsilon}{2} \quad (3)$$

pour tout $p \geq N$ et pour tout $\sigma \in S$. Soit ensuite $q \geq N$, de sorte que

$$\left| \int_q^\infty y(t, \sigma) dt \right| < \frac{\varepsilon}{2}. \quad (4)$$

La relation (2) résulte de (3) et (4), ce qui démontre la nécessité de la condition de Cauchy. Inversement, si l'inégalité (2) a lieu, alors l'intégrale (1) converge pour tout $\sigma \in S$ en vertu de 11.12.d. En passant à la limite pour $q \rightarrow \infty$ dans la relation (2) on obtient

$$\left| \int_p^\infty y(t, \sigma) dt \right| \leq \varepsilon$$

pour tous les $p \geq N$ et $\sigma \in S$, ce qui signifie la convergence uniforme de l'intégrale (1) sur l'ensemble S . Le théorème est démontré.

11.47.a. Le critère de Cauchy, à son tour, sert de base à des critères particuliers de convergence uniforme dont le suivant est un des plus simples.

Soit $f(t)$ une fonction non négative dont l'intégrale converge

$$\int_a^{\infty} f(t) dt. \quad (1)$$

Nous l'appellerons *majorant intégrable* de toute fonction $y(t)$ telle que $|y(t)| \leq f(t)$.

Théorème. Si la fonction $y(t, \sigma)$, pour tout $\sigma \in S$, possède un même majorant intégrable $f(t)$, alors l'intégrale

$$\int_a^{\infty} y(t, \sigma) dt$$

converge uniformément sur S .

La proposition résulte de l'inégalité

$$\left| \int_p^q y(t, \sigma) dt \right| \leq \int_p^q |y(t, \sigma)| dt \leq \int_p^q f(t) dt$$

en appliquant le critère de Cauchy 11.46i.

b. Exemple. Montrons que l'intégrale impropre

$$\int_0^{\infty} \frac{\cos \sigma t - \cos \beta t}{t^2} dt \quad (0 \leq \sigma \leq 1, 0 \leq \beta \leq 1) \quad (2)$$

converge uniformément par rapport au paramètre σ dans l'intervalle $0 \leq \sigma \leq 1$. Établissons préalablement les inégalités

$$\left| \frac{\cos \sigma t - \cos \beta t}{t^2} \right| \leq \begin{cases} \frac{2}{t^2} & \text{pour } t \geq 1, \\ 2 & \text{pour } t \leq 1 \end{cases}$$

pour tous les σ et β de l'intervalle $[0, 1]$.

La première inégalité est immédiate. Pour démontrer la deuxième, appliquons la formule résultant de la formule de Taylor 8.23 (4):

$$\cos \alpha t = 1 - \frac{\alpha^2 t^2}{2} + t^4 G(\alpha, t),$$

où

$$|G(\alpha, t)| = \left| \frac{\alpha^4}{4!} - \frac{\alpha^6}{6!} t^2 + \dots \right| \leq \frac{1}{4!} + \frac{1}{6!} + \dots < \frac{1}{2}$$

pour $|\alpha| \leq 1$, $|t| \leq 1$.

Il en découle, si $0 \leq \sigma \leq 1$, $0 \leq \beta \leq 1$, $0 \leq t \leq 1$, que

$$\left| \frac{\cos \sigma t - \cos \beta t}{t^2} \right| = \left| \frac{\beta^2 - \sigma^2}{2} + t^2 (G(\sigma, t) - G(\beta, t)) \right| < \frac{1}{2} + \frac{1}{2} + \frac{1}{2} < 2,$$

ce qu'il fallait démontrer.

Par conséquent, la fonction

$$f(t) = \begin{cases} \frac{2}{t^2} & \text{pour } t \geq 1 \\ 2 & \text{pour } 0 \leq t \leq 1 \end{cases}$$

représente un majorant intégrable de toutes les fonctions $\frac{\cos \sigma t - \cos \beta t}{t^2}$, $\sigma, \beta \in [0, 1]$. Donc, l'intégrale (2) converge uniformément par rapport à σ , pour $\sigma \in [0, 1]$, et représente une fonction continue de σ dans cet intervalle. L'expression explicite de cette fonction sera donnée dans 11.49.b.

11.48. Produit de convolution de deux fonctions. On appelle *produit de convolution* ou *convolution* de deux fonctions complexes $g(t)$ et $f(t)$ définies pour $-\infty < t < \infty$ l'intégrale de troisième espèce

$$h(t) = \int_{-\infty}^{\infty} f(\sigma) g(t - \sigma) d\sigma. \quad (1)$$

Les conditions d'existence de cette fonction et ses propriétés sont formulées ci-dessous.

Théorème. Si $f(t)$ et $g(t)$ sont continues, bornées et absolument intégrables sur l'axe $-\infty < t < \infty$, alors $h(t)$ existe pour tout t , est continue, bornée et absolument intégrable sur l'axe $-\infty < t < \infty$, et, de plus,

$$\int_{-\infty}^{\infty} h(t) dt = \int_{-\infty}^{\infty} f(t) dt \int_{-\infty}^{\infty} g(t) dt.$$

Démonstration. Si $|g(t)| \leq C$, alors la fonction à intégrer admet la majoration $|f(\sigma) g(t - \sigma)| \leq C |f(t)|$, d'où la convergence uniforme de l'intégrale (1) par rapport au paramètre $\sigma \in (-\infty, \infty)$. Ensuite, la fonction $h(t)$ (1) est bornée car

$$|h(t)| \leq \int_{-\infty}^{\infty} |f(\sigma)| |g(t - \sigma)| d\sigma \leq C \int_{-\infty}^{\infty} |f(\sigma)| d\sigma.$$

Montrons que la fonction sous le signe somme dans (1) est continue par rapport à l'ensemble des variables dans tout rectangle fini $a \leq t \leq b$, $\alpha \leq \sigma \leq \beta$. Soit $|g(t)| \leq C$ et $|f(t)| \leq C$. Pour un $\varepsilon > 0$ donné, trouvons un $\delta > 0$ tel que $|\sigma' - \sigma''| < \delta$, $\sigma', \sigma'' \in [\alpha, \beta]$ implique $|f(\sigma') - f(\sigma'')| < \varepsilon$ et $|t' - t''| < 2\delta$, $t', t'' \in [a - \beta, b - \alpha]$ implique $|g(t') - g(t'')| < \varepsilon$; lorsque

$|\sigma' - \sigma''| < \delta$, $|t' - t''| < \delta$, nous avons alors

$$|(t' - \sigma') - (t'' - \sigma'')| < 2\delta$$

et, pour $\sigma', \sigma'' \in [\alpha, \beta]$, $t', t'' \in [a, b]$, nous avons

$$|f(\sigma')g(t' - \sigma') - f(\sigma'')g(t'' - \sigma'')| \leq |f(\sigma') - f(\sigma'')||g(t' - \sigma')| + \\ + |f(\sigma'')||g(t' - \sigma') - g(t'' - \sigma'')| \leq C\varepsilon + C\varepsilon = 2C\varepsilon,$$

ce qui signifie que la fonction $f(\sigma)g(t - \sigma)$ est continue par rapport à l'ensemble des variables dans le rectangle mentionné. En vertu du théorème 11.43, la fonction $h(t)$ est continue en t .

A présent, vérifions l'intégrabilité absolue de $h(t)$. Notons d'abord que les fonctions $|f(\sigma)|$ et $|g(t)|$ satisfont aux mêmes conditions que $f(\sigma)$ et $g(t)$; donc, l'intégrale

$$\int_{-\infty}^{\infty} |f(\sigma)||g(t - \sigma)| d\sigma$$

converge uniformément par rapport à t elle aussi, la fonction sous le signe somme étant continue par rapport à l'ensemble t et σ dans tout rectangle $a \leq t \leq b$, $\alpha \leq \sigma \leq \beta$. En appliquant à cette intégrale le théorème 11.44 on trouve

$$\begin{aligned} \int_{-\tau}^{\tau} |h(t)| dt &\leq \int_{-\tau}^{\tau} \left\{ \int_{-\infty}^{\infty} |f(\sigma)||g(t - \sigma)| d\sigma \right\} dt = \\ &= \int_{-\infty}^{\infty} |f(\sigma)| \left\{ \int_{-\tau}^{\tau} |g(t - \sigma)| dt \right\} d\sigma = \int_{-\infty}^{\infty} |f(\sigma)| \left\{ \int_{-\tau - \sigma}^{\tau - \sigma} |g(t)| dt \right\} d\sigma \leq \\ &= \int_{-\infty}^{\infty} |f(\sigma)| \left\{ \int_{-\infty}^{\infty} |g(t)| dt \right\} d\sigma = \int_{-\infty}^{\infty} |g(t)| dt \int_{-\infty}^{\infty} |f(\sigma)| d\sigma, \end{aligned}$$

d'où l'existence de l'intégrale $\int_{-\infty}^{\infty} |h(t)| dt$. Enfin, toujours d'après le théorème 11.44, on a

$$\begin{aligned} \int_{-\tau}^{\tau} h(t) dt &= \int_{-\tau}^{\tau} \left\{ \int_{-\infty}^{\infty} f(\sigma)g(t - \sigma) d\sigma \right\} dt = \\ &= \int_{-\infty}^{\infty} f(\sigma) \left\{ \int_{-\tau}^{\tau} g(t - \sigma) dt \right\} d\sigma = \int_{-\infty}^{\infty} f(\sigma) \left\{ \int_{-\tau - \sigma}^{\tau - \sigma} g(t) dt \right\} d\sigma = \\ &= \int_{-\infty}^{\infty} f(\sigma) \left[\int_{-\infty}^{\infty} g(t) dt - \int_{|t - \sigma| > \tau} g(t) dt \right] d\sigma = \\ &= \int_{-\infty}^{\infty} g(t) dt \int_{-\infty}^{\infty} f(\sigma) d\sigma - \int_{-\infty}^{\infty} f(\sigma) \left\{ \int_{|t - \sigma| > \tau} g(t) dt \right\} d\sigma. \end{aligned}$$

Démontrons que la dernière intégrale tend vers 0 pour $\tau \rightarrow \infty$, ce qui prouvera la dernière conclusion du théorème. Pour un $\varepsilon > 0$ donné trouvons un ρ de façon à avoir

$$\int_{|\sigma| \geq \rho} |f(\sigma)| d\sigma < \varepsilon$$

et un τ tel que

$$\int_{|t| > \tau - \rho} |g(t)| dt < \varepsilon.$$

Ceci fait, on a

$$\begin{aligned} & \left| \int_{-\infty}^{\infty} f(\sigma) \left\{ \int_{|t-\sigma| \geq \tau} g(t) dt \right\} d\sigma \right| \leq \\ & \leq \left| \int_{|\sigma| \geq \rho} f(\sigma) \left\{ \int_{|t-\sigma| \geq \tau} g(t) dt \right\} d\sigma \right| + \left| \int_{|\sigma| \leq \rho} f(\sigma) \left\{ \int_{|t-\sigma| \geq \tau} g(t) dt \right\} d\sigma \right| \leq \\ & \leq \int_{|\sigma| \geq \rho} |f(\sigma)| \left\{ \int_{-\infty}^{\infty} |g(t)| dt \right\} d\sigma + \int_{|\sigma| \leq \rho} |f(\sigma)| \left\{ \int_{|t| \geq \tau - \rho} |g(t)| dt \right\} d\sigma \leq \\ & \leq \varepsilon \left\{ \int_{-\infty}^{\infty} |g(t)| dt + \int_{-\infty}^{\infty} |f(\sigma)| d\sigma \right\}, \end{aligned}$$

d'où la conclusion cherchée. Le théorème est démontré.

L'opération de convolution de deux fonctions $f(t)$ et $g(t)$ est désignée par le symbole $*$. Nous avons démontré que

$$\int_{-\infty}^{\infty} |f(t) * g(t)| dt = \int_{-\infty}^{\infty} f(t) dt \cdot \int_{-\infty}^{\infty} g(t) dt.$$

Remarque. Si dans une intégrale propre

$$h_{\tau}(t) = \int_{-\tau}^{\tau} f(\sigma) g(t - \sigma) d\sigma$$

on fait la substitution $t - \sigma = \eta$, $\sigma = t - \eta$, on obtient

$$h_{\tau}(t) = \int_{t-\tau}^{t+\tau} f(t - \eta) g(\eta) d\eta.$$

En passant à la limite pour $\tau \rightarrow \infty$ on trouve

$$\begin{aligned} f(t) * g(t) &= h(t) = \int_{-\infty}^{\infty} f(\sigma) g(t - \sigma) d\sigma = \\ &= \int_{-\infty}^{\infty} f(t - \eta) g(\eta) d\eta = g(t) * f(t), \end{aligned}$$

de sorte que la convolution est symétrique par rapport aux fonctions $f(t)$ et $g(t)$.

11.49. Critère d'Abel-Dirichlet de convergence uniforme des intégrales.

a. Il s'applique à certains cas où le critère 11.47.a ne convient pas. Soit une intégrale

$$\int_a^{\infty} u(t) v(t, \sigma) dt, \quad \sigma \in S. \quad (1)$$

Si la fonction (complexe) $u(t)$ possède sa dérivée continue par morceaux absolument intégrable, tend vers zéro pour $t \rightarrow \infty$ et que

$$\left| \int_a^T v(t, \sigma) dt \right| \leq C, \quad \text{où } C \text{ ne dépend pas de } T \text{ ni de } \sigma \in S, \text{ alors}$$

l'intégrale (1) converge uniformément sur S .

En passant à la démonstration, posons $V(t, \sigma) = \int_a^t v(\theta, \sigma) d\theta$ et intégrons par parties :

$$\int_T^{T'} u(t) v(t, \sigma) dt = u(t) V(t, \sigma) \Big|_T^{T'} - \int_T^{T'} u'(t) V(t, \sigma) dt.$$

Ici nous avons, comme dans 11.17. a :

$$\begin{aligned} |u(T') V(T', \sigma) - u(T) V(T, \sigma)| &\leq 2C \max \{u(T), u(T')\}; \\ \left| \int_T^{T'} u(t) v(t, \sigma) dt \right| &\leq 2Cu(T) + C \int_T^{T'} u'(t) dt = 3Cu(T). \end{aligned}$$

Vu l'hypothèse du théorème, les quantités obtenues tendent vers 0 pour $T \rightarrow \infty$, ce qu'il fallait démontrer.

b. Exemple. Nous avons calculé l'intégrale $\int_0^{\infty} \frac{\sin t\sigma}{t} dt = \frac{\pi}{2} \quad (\sigma > 0)$ dans 11.33. Montrons qu'elle converge uniformément dans tout intervalle $0 < \alpha \leq \sigma \leq \beta$. En effet, lorsque $t \rightarrow \infty$, la fonction $\frac{1}{t}$ tend vers 0 et possède sa dérivée continue intégrable $-1/t^2$, et l'on a

$$\left| \int_a^T \sin t\sigma dt \right| = \left| \frac{\cos T\sigma - \cos a\sigma}{\sigma} \right| \leq \frac{2}{\sigma} \leq \frac{2}{\alpha}.$$

de sorte que l'on peut appliquer le critère d'Abel-Dirichlet. En intégrant par rapport au paramètre $\sigma \in [\alpha, \beta]$ et en se servant de 11.44 on obtient la valeur de l'intégrale considérée dans 11.47 :

$$\begin{aligned} \int_0^{\infty} \frac{\cos \alpha t - \cos \beta t}{t^2} dt &= \int_0^{\infty} \left\{ \int_{\alpha}^{\beta} \frac{\sin t \sigma}{t} d\sigma \right\} dt = \\ &= \int_{\alpha}^{\beta} \left\{ \int_0^{\infty} \frac{\sin t \sigma}{t} dt \right\} d\sigma = \frac{\pi}{2} (\beta - \alpha). \end{aligned} \quad (2)$$

La convergence uniforme de cette intégrale, pour $\alpha \in [0, 1]$, étant démontrée, on peut passer à la limite dans (2) pour $\alpha \rightarrow 0$. On trouvera la valeur d'une nouvelle intégrale impropre :

$$\int_0^{\infty} \frac{1 - \cos \beta t}{t^2} dt = 2 \int_0^{\infty} \frac{\sin^2 \frac{\beta t}{2}}{t^2} dt = \frac{\pi \beta}{2}. \quad (3)$$

§ 11.5. Fonctions gamma et bêta d'Euler

11.51. *Fonction gamma d'Euler* $\Gamma(\tau)$ est définie par la formule

$$\Gamma(\tau) = \int_0^{\infty} t^{\tau-1} e^{-t} dt. \quad (1)$$

Cette intégrale impropre de troisième espèce représente la somme de l'intégrale de première espèce

$$\int_1^{\infty} t^{\tau-1} e^{-t} dt$$

et de l'intégrale de deuxième espèce

$$\int_0^1 t^{\tau-1} e^{-t} dt;$$

la première converge pour tout τ réel, la seconde pour tout $\tau > 0$. Par conséquent, l'égalité (1) définit $\Gamma(\tau)$ pour tout $\tau > 0$. Si $\beta \geq \tau \geq \alpha > 0$, la fonction $t^{\tau-1} e^{-t}$ admet le majorant intégrable $\max \{t^{\alpha-1} e^{-t}, t^{\beta-1} e^{-t}\}$. Donc, l'intégrale (1) converge uniformément sur n'importe quel intervalle $[\alpha, \beta]$ et représente, d'après le théorème 11.43, une fonction continue de τ . Ainsi, $\Gamma(\tau)$ est une fonction continue de τ sur $(0, \infty)$. La dérivée par rapport à τ de la fonction à intégrer a la forme $t^{\tau-1} \cdot \ln t \cdot e^{-t}$ et possède le majorant intégrable $t^{\beta-1} \ln t \cdot e^{-t}$ sur tout intervalle $0 < \alpha \leq \tau \leq \beta$; or la fonction

$t^{\tau-1} \ln t \cdot e^{-t}$ est continue par rapport à $t \in [0, \infty]$, pour tout $\tau > 0$, et par rapport à t et τ dans tout rectangle $0 \leq t \leq b$, $\alpha \leq \tau \leq \beta$. La fonction $\Gamma(\tau)$ a donc, d'après le théorème 11.45.a, la dérivée

$$\Gamma'(\tau) = \int_0^{\infty} t^{\tau-1} \ln t \cdot e^{-t} dt \quad (\tau > 0),$$

qui est continue pour la même raison que la fonction $\Gamma(\tau)$ elle-même. On peut continuer la dérivation en appliquant chaque fois le théorème 11.45.a; par conséquent, $\Gamma(\tau)$ a les dérivées de tous les ordres.

11.52. En intégrant 11.51 (1) par parties on obtient

$$\Gamma(\tau) = \int_0^{\infty} t^{\tau-1} e^{-t} dt = \frac{t^{\tau}}{\tau} e^{-t} \Big|_{t=0}^{\infty} + \frac{1}{\tau} \int_0^{\infty} t^{\tau} e^{-t} dt,$$

ou bien

$$\Gamma(\tau+1) = \tau \Gamma(\tau). \quad (1)$$

L'égalité (1) représente la *relation de récurrence de la fonction gamma*. En l'appliquant plusieurs fois on voit que

$$\Gamma(\tau+n) = (\tau+n-1)(\tau+n-2) \dots (\tau+1) \tau \Gamma(\tau); \quad (2)$$

par conséquent, si l'on connaît les valeurs de la fonction gamma dans un intervalle de longueur 1, on peut trouver ses valeurs en tous les autres points du demi-axe $\tau > 0$. Etant donné

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = 1,$$

on obtient

$$\Gamma(n+1) = n(n-1) \dots 1 = n!.$$

Il en résulte que la fonction gamma représente un prolongement de la fonction $n!$ définie pour les valeurs naturelles de n seulement sur l'ensemble de tous les nombres positifs. Ensuite, la continuité de $\Gamma(\tau)$ pour $\tau = 1$ et la relation (1) ont pour conséquence:

$$\Gamma(\tau) = \frac{\Gamma(\tau+1)}{\tau} \rightarrow \infty$$

lorsque $\tau \rightarrow 0$.

11.53. *Fonction bêta et sa relation avec la fonction gamma*. L'intégrale

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx \quad (1)$$

fonction de deux paramètres p et q s'appelle *fonction bêta d'Euler*. Elle existe pour toutes les valeurs positives de p et q : en tant qu'intégrale propre si $p, q \geq 1$ et impropre pour toutes les autres valeurs de p et q . La substitution $x = \frac{\theta}{1+\theta}$ met l'intégrale (1) sous la forme

$$B(p, q) = \int_0^{\infty} \frac{\theta^{p-1} d\theta}{(1+\theta)^{p+q}}. \quad (2)$$

Montrons qu'il est possible d'exprimer la fonction bêta par la fonction gamma.

En faisant, dans l'expression de la fonction gamma

$$\Gamma(\tau) = \int_0^{\infty} t^{\tau-1} e^{-t} dt,$$

la substitution $t = \theta y$ on obtient

$$\frac{\Gamma(\tau)}{\theta^{\tau}} = \int_0^{\infty} y^{\tau-1} e^{-\theta y} dy.$$

Remplaçons ici τ par $p+q$ et θ par $1+\theta$:

$$\frac{\Gamma(p+q)}{(1+\theta)^{p+q}} = \int_0^{\infty} y^{p+q-1} e^{-y} e^{-\theta y} dy.$$

Multiplions les deux membres de cette égalité par θ^{p-1} et intégrons de 0 à n :

$$\Gamma(p+q) \int_0^n \frac{\theta^{p-1} d\theta}{(1+\theta)^{p+q}} = \int_0^n \left\{ \int_0^{\infty} \theta^{p-1} y^{p+q-1} e^{-y} e^{-\theta y} dy \right\} d\theta. \quad (3)$$

Lorsque $n \rightarrow \infty$, le premier membre tend vers $\Gamma(p+q) B(p, q)$.

La fonction sous le signe somme dans le deuxième membre a le majorant intégrable $n^{p-1} y^{p+q-1} e^{-y}$, donc, d'après le théorème 11.44, on peut permuter les intégrations:

$$\begin{aligned} & \int_0^n \left\{ \int_0^{\infty} \theta^{p-1} y^{p+q-1} e^{-y} e^{-\theta y} dy \right\} d\theta = \\ &= \int_0^{\infty} y^{p+q-1} e^{-y} \left\{ \int_0^n \theta^{p-1} e^{-\theta y} d\theta \right\} dy = \\ &= \int_0^{\infty} y^{p+q-1} e^{-y} \left\{ \frac{1}{y^p} \int_0^{ny} t^{p-1} e^{-t} dt \right\} dy = \int_0^{\infty} y^{q-1} e^{-y} F_n(y) dy, \end{aligned}$$

où la fonction

$$F_n(y) = \int_0^{ny} t^{p-1} e^{-t} dt,$$

pour $n \rightarrow \infty$, tend vers la fonction $F(y)$ qui vaut 0 pour $y = 0$ et $\Gamma(p)$ pour $y > 0$. Cette convergence n'est pas uniforme sur l'intervalle $0 \leq y \leq b$ (ce qui résulte déjà de la discontinuité de la fonction limite). Au contraire, la convergence $F_n(y) \rightarrow F(y)$ est uniforme sur tout intervalle $h \leq y \leq b$ avec $h > 0$, car

$$0 < \Gamma(p) - \int_0^{ny} t^{p-1} e^{-t} dt \leq \Gamma(p) - \int_0^{nh} t^{p-1} e^{-t} dt.$$

En outre, vu que la fonction $F_n(y)$ tend en croissant vers $F(y)$, la famille des fonctions $y^{q-1} e^{-y} F_n(y)$ admet le majorant intégrable $y^{q-1} e^{-y} \Gamma(p)$; donc, en vertu du théorème 11.43, on a

$$\lim_{n \rightarrow \infty} \int_h^\infty y^{q-1} e^{-y} F_n(y) dy = \Gamma(p) \int_h^\infty y^{q-1} e^{-y} dy.$$

Or, pour tout $h > 0$,

$$0 \leq \int_0^h y^{q-1} e^{-y} F_n(y) dy \leq \Gamma(p) \int_0^h y^{q-1} dy = \Gamma(p) \frac{h^q}{q}. \quad (4)$$

Etant donné un $\varepsilon > 0$, trouvons un $h > 0$ de façon à avoir

$$\Gamma(p) \frac{h^q}{q} < \frac{\varepsilon}{3}. \quad (5)$$

Ensuite, h étant choisi, trouvons un N tel que

$$0 < \Gamma(p) \int_h^\infty y^{q-1} e^{-y} dy - \int_h^\infty y^{q-1} e^{-y} F_n(y) dy < \frac{\varepsilon}{3} \quad (6)$$

pour $n > N$. Notons enfin que, pour h choisi,

$$\begin{aligned} 0 &< \Gamma(p) \Gamma(q) - \Gamma(p) \int_h^\infty y^{q-1} e^{-y} dy = \\ &= \Gamma(p) \int_0^h y^{q-1} e^{-y} dy < \Gamma(p) \frac{h^q}{q} < \frac{\varepsilon}{3}. \end{aligned} \quad (7)$$

Il résulte des inégalités (4), (5), (6), (7) que

$$0 < \Gamma(p) \Gamma(q) - \int_0^\infty y^{q-1} e^{-y} F_n(y) dy < \varepsilon,$$

d'où

$$\lim_{n \rightarrow \infty} \int_0^{\infty} y^{q-1} e^{-y} F_n(y) dy = \Gamma(p) \Gamma(q).$$

En passant maintenant à la limite dans (3), on obtient

$$B(p, q) = \frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)}. \quad (8)$$

C'est l'expression cherchée de la fonction bêta par la fonction gamma.

11.54. Plusieurs intégrales trigonométriques s'expriment, à leur tour, par la fonction bêta. Ainsi, en faisant la substitution $\sin^2 \theta = x$ dans l'intégrale

$$I = \int_0^{\pi/2} \sin^{p-1} \theta \cos^{q-1} \theta d\theta$$

on a

$$\begin{aligned} I &= \int_0^1 x^{(p-1)/2} (1-x)^{(q-1)/2} \frac{dx}{2\sqrt{x(1-x)}} = \\ &= \frac{1}{2} \int_0^1 x^{\frac{p}{2}-1} (1-x)^{\frac{q}{2}-1} dx = \frac{1}{2} B\left(\frac{p}{2}, \frac{q}{2}\right). \end{aligned}$$

11.55. Formule des compléments. En posant $q = 1 - p$ on tire de la formule 11.53 (8):

$$\Gamma(p) \Gamma(1-p) = B(p, 1-p) = \int_0^{\infty} \frac{\theta^{p-1} d\theta}{1+\theta}. \quad (1)$$

Pour calculer cette intégrale, considérons la fonction analytique

$$w(z) = \frac{z^{p-1}}{1+z}$$

dans le plan de la variable complexe z coupé le long du demi-axe réel positif. La fonction est définie de façon unique si l'on pose, sur le bord supérieur de la coupure,

$$w(x+i0) = \frac{x^{p-1}}{1+x}, \quad (2)$$

où x^{p-1} est considéré au sens de 5.54. Sur le bord inférieur de la coupure, on a, après avoir contourné l'origine des coordonnées dans le sens positif, $z = x - i0 = xe^{2\pi i}$,

$$w(x-i0) = \frac{x^{p-1}}{1+x} e^{2\pi i(p-1)}. \quad (3)$$

Considérons le contour fermé L (fig. 11.7) formé du segment $[0, R]$ ($R > 1$) de l'axe des x sur le bord supérieur de la coupure, de la circonférence C_R de rayon R centrée en l'origine des coordonnées

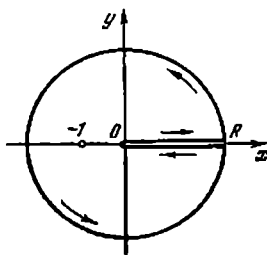


Fig. 11.7.

et du segment $[R, 0]$ sur le bord inférieur de la coupure. A l'intérieur du contour L , il n'y a qu'un seul point singulier de la fonction $w(z)$, qui est le pôle du premier ordre en $z = -1$. D'après 10.42 a, on a

$$\int_L w(z) dz = 2\pi i \cdot \text{Res } w(z)|_{z=-1} = 2\pi i e^{i\pi(p-1)}. \quad (4)$$

Or,

$$\int_L w(z) dz = \int_0^R w(x+i0) dx + \int_{C_R} w(z) dz + \int_R^0 w(x-i0) dx. \quad (5)$$

Lorsque $|z|=R$, on a $|w(z)| \leq \frac{R^{p-1}}{R-1} \leq cR^{p-2}$, d'où, pour $p < 1$,

$$\left| \int_{C_R} w(z) dz \right| \leq cR^{p-2} 2\pi R = 2\pi c R^{p-1} \rightarrow 0. \quad (6)$$

En passant à la limite pour $R \rightarrow \infty$ dans (5) et compte tenu de (2)-(5), on trouve

$$2\pi i e^{i\pi(p-1)} = \int_0^\infty \frac{x^{p-1}}{1+x} dx \cdot (1 - e^{2\pi i(p-1)})$$

donc

$$\begin{aligned} \int_0^\infty \frac{x^{p-1}}{1+x} dx &= 2\pi i \frac{e^{i\pi(p-1)}}{1 - e^{2\pi i(p-1)}} = \frac{2\pi i}{e^{-i\pi(p-1)} - e^{i\pi(p-1)}} = \\ &= -\frac{\pi}{\sin \pi(p-1)} = \frac{\pi}{\sin \pi p}. \end{aligned}$$

Définitivement, on obtient

$$\Gamma(p) \Gamma(1-p) = B(p, 1-p) = \frac{\pi}{\sin p\pi}. \quad (7)$$

C'est la formule des compléments. En particulier, en y posant $p = 1/2$, on a

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \quad (8)$$

En se servant de la relation de récurrence on obtient

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1}{2} \sqrt{\pi},$$

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \sqrt{\pi}, \dots, \Gamma\left(\frac{2n+1}{2}\right) = \frac{(2n-1)!!}{2^n} \sqrt{\pi}.$$

Le graphique de la fonction gamma pour $0 < p \leq 5$ est représenté fig. 11.8.

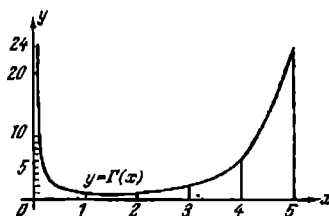


Fig. 11.8.

11.56. A l'aide de la fonction gamma, on peut calculer les intégrales essentielles pour la théorie des probabilités:

$$I_m = \int_0^{\infty} x^m e^{-ax^2} dx \quad (a > 0, m > -1). \quad (1)$$

Notamment, la substitution $ax^2 = t$, $x = \sqrt{\frac{t}{a}}$, $dx = \frac{1}{2} \frac{dt}{\sqrt{at}}$ met les intégrales I_m sous la forme

$$I_m = \frac{1}{2} \cdot \frac{2}{a^{(m+1)/2}} \int_0^{\infty} t^{(m-1)/2} e^{-t} dt = \frac{1}{2a^{(m+1)/2}} \Gamma\left(\frac{m+1}{2}\right).$$

En particulier, pour $a = 1$, on a

$$I_0 = \int_0^{\infty} e^{-x^2} dx = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1}{2} \sqrt{\pi}. \quad (2)$$

11.57. Expression asymptotique de la fonction gamma. Pour les grandes valeurs de τ , la fonction gamma

$$\Gamma(\tau) = \int_0^{\infty} t^{\tau-1} e^{-t} dt$$

admet une simple représentation asymptotique.

a. L e m m e. Supposons qu'une fonction $f(x) \geq 0$ définie pour $x > 0$ vaille 0 pour $x = 1$, décroisse pour $0 < x < 1$, croisse pour $x > 1$, admette la représentation

$$f(x) = a(x-1)^2 + (x-1)^{2+2\delta} \psi(x),$$

$$\delta > 0, |\psi(x)| \leq M, a > 0,$$

au voisinage du point $x = 1$ et vérifie la condition

$$|f(x)| \geq bx, \quad b > 0, \quad (1)$$

pour $x \geq c > 1$. Alors

$$I(s) = \int_0^{\infty} e^{-sf(x)} dx \quad (2)$$

est un infiniment petit équivalent à $\sqrt{\frac{\pi}{sa}}$ lorsque $s \rightarrow \infty$.

D é m o n s t r a t i o n. La convergence de l'intégrale (2) pour $s > 0$ résulte de l'estimation (1). Pour $0 < \varepsilon < 1$, on a

$$I(s) = \left\{ \int_0^{1-\varepsilon} + \int_{1-\varepsilon}^{1+\varepsilon} + \int_{1+\varepsilon}^c + \int_c^{\infty} \right\} e^{-sf(x)} dx.$$

Ensuite, si ε est assez petit, on a par hypothèse

$$\int_0^{1-\varepsilon} e^{-sf(x)} dx < e^{-sf(1-\varepsilon)} < e^{-s \frac{a}{2} \varepsilon^2}, \quad (3)$$

$$\int_{1+\varepsilon}^c e^{-sf(x)} dx < (c-1) e^{-sf(1+\varepsilon)} < (c-1) e^{-s \frac{a}{2} \varepsilon^2}, \quad (4)$$

$$\int_c^{\infty} e^{-sf(x)} dx < \int_c^{\infty} e^{-sbx} dx = \frac{e^{-sbc}}{sb}. \quad (5)$$

En passant au calcul de l'intégrale de $1 - \varepsilon$ à $1 + \varepsilon$ on voit que, dans cet intervalle,

$$\begin{aligned} & |e^{-s(x-1)^{2+2\delta}\psi(x)} - 1| = \\ & = \left| s(x-1)^{2+2\delta}\psi(x) + \frac{1}{2}s^2(x-1)^{2(2+2\delta)}\psi^2(x) + \dots \right| \leq \\ & \leq Mse^{2+2\delta} + \frac{1}{2}M^2s^2e^{2(2+2\delta)} + \dots \leq 2Mse^{2+2\delta}. \end{aligned}$$

dès que

$$Mse^{2+2\delta} \leq 1, \quad (6)$$

Le nombre ε est, pour le moment, arbitraire; si $s \rightarrow \infty$, posons

$$\varepsilon = s^{-\frac{1}{2+\delta}}.$$

Alors

$$se^{2+2\delta} = s^{1-\frac{2+2\delta}{2+\delta}} = s^{-\frac{\delta}{2+\delta}} \rightarrow 0 \quad (s \rightarrow \infty),$$

et la condition (6) est bien satisfaite. D'autre part, vu le choix de ε , on a

$$s\varepsilon^2 = s^{1-\frac{2}{2+\delta}} = s^{\frac{\delta}{2+\delta}} \rightarrow \infty \quad (s \rightarrow \infty).$$

A présent, pour $s \rightarrow \infty$ et n'importe quel $x \in (1 - \varepsilon, 1 + \varepsilon)$ on a par construction

$$|e^{-s(x-1)^{2+2\delta}\psi(x)} - 1| \leq 2Mse^{2+2\delta} \rightarrow 0,$$

de sorte que

$$e^{-s(x-1)^{2+2\delta}\psi(x)} = 1 + o(1),$$

où $o(1) \rightarrow 0$ lorsque $s \rightarrow \infty$. Par conséquent,

$$\begin{aligned} & \int_{1-\varepsilon}^{1+\varepsilon} e^{-sf(x)} dx = \int_{1-\varepsilon}^{1+\varepsilon} e^{-sa(x-1)^2(1+o(1))} dx = \\ & = [1+o(1)] \int_{1-\varepsilon}^{1+\varepsilon} e^{-sa(x-1)^2} dx = [1+o(1)] \frac{1}{\sqrt{sa}} \int_{-\varepsilon\sqrt{sa}}^{\varepsilon\sqrt{sa}} e^{-u^2} du. \end{aligned}$$

Vu que $\varepsilon\sqrt{s} \rightarrow \infty$, la dernière intégrale tend vers la limite

$$\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$$

lorsque $s \rightarrow \infty$ (11.56), de sorte que

$$\int_{-e\sqrt{sa}}^{e\sqrt{sa}} e^{-u^2} du = \sqrt{\pi} (1 + o(1)),$$

où $o(1) \rightarrow 0$ pour $s \rightarrow \infty$. Donc

$$\int_{1-e}^{1+e} e^{-s/(x)} dx = (1 + o(1)) \sqrt{\frac{\pi}{sa}}.$$

Quant aux autres termes de (3), (4), (5), ils tendent tous, en vertu de $se^2 = s^{5/(2+\delta)}$, vers 0 pour $s \rightarrow \infty$, selon la loi exponentielle. Définitivement,

$$I(s) = \int_0^\infty e^{-s/(x)} dx = (1 + o(1)) \sqrt{\frac{\pi}{sa}} + o\left(\frac{1}{\sqrt{s}}\right) = \sqrt{\frac{\pi}{sa}} [1 + o(1)].$$

b. Mettons à présent l'expression de la fonction gamma sous une telle forme que l'on puisse appliquer le lemme. Avec la substitution $t = sx$, on a

$$\begin{aligned} \Gamma(s+1) &= \int_0^\infty t^s e^{-t} dt = s^{s+1} \int_0^\infty x^s e^{-sx} dx = \\ &= s^{s+1} \int_0^\infty e^{-sx + s \ln x} dx = s^{s+1} e^{-s} \int_0^\infty e^{-s(x - \ln x - 1)} dx = \\ &= s^{s+1} e^{-s} \int_0^\infty e^{-s f(x)} dx, \end{aligned}$$

où

$$f(x) = x - \ln x - 1.$$

Il est aisé de voir que la fonction $f(x)$ vérifie l'hypothèse du lemme. De plus, $f(1) = f'(1) = 0$, $f''(1) = 1$, $f'''(1) = -2$; au voisinage du point $x = 1$, on a

$$f(x) = \frac{(x-1)^2}{2} f''(1) - \frac{E}{3} (x-1)^3, \quad \lim_{x \rightarrow 1} E = 1,$$

de sorte que la constante a vaut $1/2$. En appliquant le lemme on trouve

$$\Gamma(s+1) = s^{s+1} e^{-s} \sqrt{\frac{2\pi}{s}} [1 + o(1)] \approx s^{s+\frac{1}{2}} e^{-s} \sqrt{2\pi},$$

ce qui représente l'expression asymptotique cherchée.

En particulier, si $s = n$ est un nombre naturel, alors $\Gamma(n+1) = n!$ et l'on obtient

$$n! \approx n^{n+1/2} e^{-n} \sqrt{2\pi}$$

(formule de Stirling, 1730).

11.58. Fonction gamma dans le domaine complexe. La formule

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt = \int_0^{\infty} e^{(z-1) \ln t} e^{-t} dt \quad (1)$$

qui définit la fonction gamma est valable non seulement pour les valeurs réelles $z > 0$, mais aussi pour certains z complexes. Notamment, si $z = x + iy$, $x > 0$, alors l'intégrale (1) converge, puisque la fonction à intégrer

$$e^{(x+iy-1) \ln t} e^{-t} = e^{(x-1) \ln t} e^{iy \ln t} e^{-t} = t^{x-1} e^{-t} e^{iy \ln t}$$

ne diffère de la fonction $t^{x-1} e^{-t}$ que par le facteur $e^{iy \ln t}$ de module unité. Par conséquent, la formule (1) permet de définir directement $\Gamma(z)$ pour tous les $z = x + iy$ avec $x > 0$, c'est-à-dire dans tout le demi-plan (ouvert) à droite G du plan des z . L'intégrale (1) converge uniformément à l'intérieur de G car, pour tout compact $Q \subset G$, la quantité $d = \inf_{z \in Q} x$ est positive et il y a un majorant intégrable $t^{d-1} e^{-t}$. D'après 11.45.b, la fonction $\Gamma(z)$ est analytique dans le domaine G .

Voyons s'il est possible de prolonger analytiquement la fonction $\Gamma(z)$ dans le demi-plan à gauche. Utilisons l'équation 11.52 (2)

$$\Gamma(z+n) = z(z+1) \dots (z+n-1) \Gamma(z) \quad (2)$$

Elle est établie pour les valeurs réelles de z . Or, les deux membres de l'équation étant des fonctions analytiques de z dans le domaine G , le fait qu'ils sont identiques sur l'axe réel implique leur identité dans tout le demi-plan G , en vertu du théorème d'unicité 10.39.b. En mettant (2) sous la forme

$$\Gamma(z) = \frac{\Gamma(z+n)}{z(z+1) \dots (z+n-1)} \quad (3)$$

faisons attention à ce que le second membre est défini et analytique pour tout z avec $\operatorname{Re} z > -n$, à l'exception des valeurs $z = 0, -1, \dots, -n+1$. Considérons maintenant la formule (3) comme définition de la fonction gamma. Ceci fournit un prolongement analytique de cette dernière dans le domaine $G_n = \{z: \operatorname{Re} z > -n\}$. Formellement, on a les définitions différentes pour les n différents, mais, vu l'unicité du prolongement analytique (10.39.c), toutes les

définitions donnent une même valeur de la fonction $\Gamma(z)$ quel que soit z . Comme n est arbitrairement grand, la fonction $\Gamma(z)$ se trouve définie dans tout le plan des z , à l'exception des singularités isolées aux points $z = 0, 1, \dots, -n+1, \dots$.

Il résulte de la formule (3) que toutes ces singularités sont des pôles du premier ordre. En calculant le résidu de la fonction $\Gamma(z)$ par rapport au pôle $z = -n+1$, à l'aide de la formule 10.42 (1), nous avons

$$\begin{aligned} \operatorname{Res} \Gamma(z) \Big|_{z=-n+1} &= \frac{\Gamma(z+n)}{z(z+1)\dots(z+n-2)} \Big|_{z=-n+1} = \\ &= \frac{\Gamma(1)}{(-1)(-2)\dots(-n+1)} = \frac{(-1)^{n-1}}{(n-1)!}. \end{aligned}$$

Exercices

1. Démontrer la formule de Dirichlet

$$\begin{aligned} \int_0^1 \left\{ \int_{y=0}^{1-x} x^{\lambda-1} y^{\mu-1} (1-x-y)^{\nu-1} f(x, y) dy \right\} dx = \\ = \int_0^1 \left\{ \int_{x=0}^{1-y} x^{\lambda-1} y^{\mu-1} (1-x-y)^{\nu-1} f(x, y) dx \right\} dy, \end{aligned}$$

où $0 < \lambda \leq 1$, $0 < \mu \leq 1$, $0 < \nu \leq 1$, $f(x, y)$ étant continue pour $0 \leq x \leq 1$, $0 \leq y \leq 1$.

2. Démontrer la formule de Froullani

$$\int_0^{\infty} \frac{f(ax) - f(bx)}{x} dx = f(0) \ln \frac{b}{a},$$

où $a > 0$, $b > 0$, si $f(x)$ est continue pour $0 \leq x < \infty$ et l'intégrale

$$\int_1^{\infty} \frac{f(x)}{x} dx \text{ converge.}$$

3. En se servant de l'intégration dans le plan complexe, trouver

$$I(p) = \int_0^{\infty} e^{-px^2} dx,$$

où $\operatorname{Re} p > 0$.

4. Calculer les intégrales de Fresnel

$$F_1 = \int_0^{\infty} \sin(x^2) dx, \quad F_2 = \int_0^{\infty} \cos(x^2) dx.$$

5. Calculer l'intégrale spéciale de Fourier

$$I = \int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} \frac{\sin x}{x} dx,$$

où $P(x)$ et $Q(x)$ sont des polynômes, le degré de $P(x)$ ne dépasse pas celui de $Q(x)$ et $Q(x)$ ne s'annule pour aucun x réel.

6. Montrer que

$$\int_0^1 \left\{ \int_1^{\infty} \frac{y-x}{(x+y)^3} dx \right\} dy = -\frac{1}{2}, \quad \int_1^{\infty} \left\{ \int_0^1 \frac{y-x}{(x+y)^3} dy \right\} dx = -1.$$

Est-ce qu'il y a contradiction avec le théorème 11.44?

7. Désignons par L le chemin dans le plan de la variable complexe z , formé de l'intervalle $[e, +\infty)$ de l'axe réel parcouru de droite à gauche, de la circonférence $|z| = e$ parcourue dans le sens positif et de l'intervalle $[e, +\infty)$ de l'axe réel parcouru de gauche à droite. Démontrer la formule

$$\Gamma(z) = \frac{1}{e^{2\pi iz} - 1} \int_L e^{-zs} s^{-1} ds,$$

qui fournit une représentation de la fonction gamma pour tout z complexe, à l'exception de ses pôles.

Historique

Plusieurs intégrales impropres sont calculées aux XVII^e et XVIII^e siècles, avant même que Cauchy ne donne une définition rigoureuse de la convergence d'une intégrale impropre (1821). C'est encore lui qui indique la méthode de calcul des intégrales impropres à l'aide du prolongement analytique et de la théorie des résidus. Les intégrales absolument convergentes sont dégagées par Dirichlet (1854), celles uniformément convergentes par Vallée Poussin (1892). Les fonctions gamma et bêta sont introduites et étudiées par Euler (années 1730).

Indications et réponses

Chapitre 1

1. *Indication.* On a

$$|x| = |x - y + y| \leq |x - y| + |y|,$$

d'où

$$|x| - |y| \leq |x - y|.$$

Remplacer ici x par y et y par x .

2. *Indication.* Prouver que le nombre $\sup A + \sup B$ possède toutes les propriétés de la borne supérieure de l'ensemble $A + B$.

3. *Indication.* Prouver que le nombre $\sup A \cdot \sup B$ possède toutes les propriétés de la borne supérieure de l'ensemble AB .

4. *Indication.* Procéder par analogie avec les exercices 2 et 3.

5. *Indication.* L'ensemble A est majoré; $\gamma = \sup A$.

Chapitre 2

1. *Indication.* Appliquer 2.35 — 2.36.

2. *Indication.* a) Dans chacun des intervalles en question, on peut choisir un point rationnel. b) Dans chaque boucle du huit, choisir un point de coordonnées rationnelles. c) A l'extérieur de tout intervalle $[0, \epsilon]$, il n'y a qu'un nombre fini de points de M .

3. *Indication.* Inverser la construction 2.33.

4. *Indication.* Les frères N forment un ensemble dénombrable.

5. *Indication.* Soient C une partie dénombrable de A et $D = A - C$. Alors $A = D + C$. $A + B = D + C + B$; utiliser l'équipotence de C et $C + B$.

6. *Indication.* Se servir de l'exercice 5 en choisissant pour B l'ensemble des nombres rationnels et en posant $A = I$. Le cas de l'ensemble T est analogue.

7. *Indication.* L'ensemble A est la réunion de deux ensembles: celui des suites contenant des zéros en nombre fini et celui des suites contenant chacune une infinité de zéros. Le premier est dénombrable, le second a la puissance du continu car il est en correspondance biunivoque avec l'ensemble des points d'un intervalle notés dans le système binaire (cf. 1.78, $p = 2$). Appliquer l'exercice 5.

8. *Indication.* Faire correspondre à toute suite (1) la suite formée de zéros et d'unités, avec les unités aux places $k_1, k_2, \dots$. Appliquer l'exercice 7.

9. *Indication.* Faire correspondre à toute suite (2) la suite croissante

$$k_1 = m_1, k_2 = m_1 + m_2, \dots, k_n = m_1 + \dots + m_n, \dots$$

10. *Indication.* En appliquant l'exercice 9 on peut faire correspondre à tout symbole ξ un tableau

$$\xi \sim \begin{cases} m_{11} & m_{12} & \dots & m_{1n} & \dots \\ m_{21} & m_{22} & \dots & m_{2n} & \dots \\ \dots & \dots & \dots & \dots & \dots \\ m_{p1} & m_{p2} & \dots & m_{pn} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{cases}$$

dont les éléments sont des nombres naturels. Ce tableau peut être mis sous forme de suite, comme dans 2.33. Ceci fait, on applique encore une fois l'exercice 9.

11. *Indication.* Simplifier convenablement la résolution de l'exercice 10.

12. *Indication.* Supposons que, à tout $t \in [0, 1]$ on fasse correspondre une fonction $f_t(x) \in E$. La famille $W \subset E$ de fonctions $f_t(x)$ n'épuise pas E car la fonction $\phi(x)$ dont la valeur pour tout x est différente de la valeur correspondante $f_x(x)$ n'appartient pas à W .

13. *Indication.* Généraliser la méthode de résolution de l'exercice 12 en remplaçant $[0, 1]$ par A et en définissant tout sous-ensemble $B \subset A$ à l'aide de la fonction qui vaut 1 sur B et 0 ailleurs.

Chapitre 3

1. *Indication.* Si $A = \left\{1, \frac{1}{2}, \frac{1}{3}, \dots\right\}$, alors $A' = \{0\}$, A^* est vide. Construire une itération.

2. *Indication.* Tout point limite de l'ensemble A' en est un pour A .

3. *Indication.* Il suffit de considérer $n = 1$. Tout point de l'ensemble A qui n'en est pas un point limite peut être recouvert par un intervalle aux extrémités rationnelles ne contenant aucun autre point de l'ensemble A . Appliquer l'exercice 1(a) du chapitre 2.

4. *Indication.* Un point qui n'est pas un point de condensation de l'ensemble A peut être recouvert par un intervalle aux extrémités rationnelles contenant un ensemble au plus dénombrable des points de l'ensemble A .

5. *Indication.* Marquer chaque intervalle aux extrémités rationnelles qui est contenu dans un intervalle du recouvrement. De tous les intervalles du recouvrement qui contiennent un même intervalle marqué laisser un seul.

8. *Indication.* Poser $U_1 = \bigcup_{x \in F_1} \left\{y : \rho(x, y) \leq \frac{1}{2} \rho(x, F_2)\right\}$; définir U_2

de manière analogue.

10. *Indication.* Pour $n = 3$, la construction se fait avec deux circonférences. L'espace M formé de quatre points x_1, x_2, x_3, x_4 , avec pour distances

$$\rho(x_1, x_2) = \rho(x_2, x_3) = \rho(x_3, x_4) = \rho(x_4, x_1) = 1,$$

$$\rho(x_1, x_3) = \rho(x_2, x_4) = 2$$

n'est plus isométrique à aucun sous-ensemble de l'espace euclidien R_3 .

11. *Indication.* Considérer la sphère $S_n \subset R_{n+1}$ de rayon 2 et de centre $(0, 0, \dots, 0, 1)$. Appliquer dans S_n l'espace $R_n \subset R_{n+1}$ formé des points $(\xi_1, \dots, \xi_n, 0)$ en se servant des droites passant par le point $(0, 0, \dots, 0, 2)$. Faire correspondre ce point au point $\infty \in R_n$. Choisir pour métrique r la métrique usuelle de l'espace R_{n+1} pour les points de S_n (« projection stéréographique »).

12. *Indication.* En fixant un point $a \in M$ trouver, pour tout couple de points x, y du plan euclidien R_2 , trois points A, X, Y avec les mêmes distances relatives (exercice 10). Définir la nouvelle distance $r(x, y)$ moyennant la projection stéréographique de R_2 , sur la sphère S_2 tangente à R_2 au point A .

13. *Réponse.* Deux éléments pour toute droite passant par l'origine.

Chapitre 4

1. *Indication.* Cf. l'exercice 1 du chapitre 1. Considérer la suite $x_n = (-1)^n$ ($n = 1, 2, \dots$).

2. *Indication.* Les limites $\lim$ et $\limsup$ sont des limites de certaines suites.

3. *Réponse.* Oui.

4. *Indication.* Si $a_{n_k} \rightarrow \overline{\lim} a_n < \overline{\lim} a_n$, poser, par exemple,

$$b_n = \begin{cases} 0 & \text{pour } n = n_k, \quad k = 1, 2, \dots, \\ \overline{\lim} a_n - \overline{\lim} a_n & \text{pour } n \neq n_k, \end{cases}$$

6. *Réponse.* $\lim x_n = \frac{1}{3}a + \frac{2}{3}b$.

7. *Indication.* Pour la suite $e_n = \frac{x_n}{\sqrt{a}} - 1$, établir une équation et l'esti-

$$\text{matlon } e_{n+1} = \frac{e_n^2}{2(1+e_n)} < \frac{e_n}{2}.$$

8. *Indication.* $v_n > v_{n+1} > x_{n+1} > x_n$.

9. *Indication.* Soit $p_1 = \dots = p_r > p_{r+1} \geq \dots \geq p_m$; alors $\lim_{n \rightarrow \infty} \left(\frac{p_k}{p_1} \right)^n = 0$

pour $k > r$.

11. *Indication.* Se servir de la définition de la limite suivant la direction $x \rightarrow a$.

12. Les réponses aux deux questions sont, en général, négatives. Exemples: soit $X = \{1, 2, \dots, n, \dots\}$, $S = \{n \rightarrow \infty\}$, $Y = \{0, 1, 1/2, 1/3, \dots\}$, $Z = \{0, 1\}$: soit, ensuite, $y_1(2n) = 1/2n$, $y_1(2n+1) = 0$, $y_2(n) = 0$ ($n = 1, 2, \dots$), $z(0) = 1$, $z(y) = 0$ pour $y \neq 0$. Alors $z(y_1(n))$ n'a pas de limite pour $n \rightarrow \infty$, et $z(y_2(n))$ possède la limite qui vaut $1 \neq \lim_{y \rightarrow 0} z(y)$.

13. *Indication.* a) et b) sont immédiats, c) résulte de 4.16.c.

Chapitre 5

1. *Réponse.* Non.

2. *Réponse.* Non.

3. *Indication.* Appliquer l'exercice 1. du chapitre 1.

4. *Réponse.* Les deux fonctions sont discontinues pour les x entiers et continues pour tous les autres x .

7. *Indication.* $\min(\xi_1, \dots, \xi_n) \leq \frac{1}{n} \sum_{k=1}^n \xi_k \leq \max(\xi_1, \dots, \xi_n)$.

8. *Indication.* Utiliser l'existence de $f(a+0)$ et $f(b-0)$.

9. *Indication.* Poser $x = \log_a t$ et appliquer 5.41.

10. *Réponse.* $]-\infty, -1-\delta]$; $]-1+\delta, 1-\delta]$; $[1+\delta, \infty]$; $\delta > 0$ est quelconque.

11. *Indication.* Considérer les couples de fonctions $\sin ax, \cos ax$; $a^x \sin x, a^x \cos x$.

12. *Indication.* En tant qu'exemple, considérer la fonction

$$y = \begin{cases} \sin \frac{1}{x} & \text{pour } x \neq 0, \\ 0 & \text{pour } x = 0. \end{cases}$$

13. *Indication.* Appliquer les formules 5.61 (1)-(3).

14. *Indication.* Tout point c tel que $|\lim_{t \rightarrow c} x(t) - x(c)| \geq q > 0$ a un voisinage dont aucun point ne possède la même propriété.

15. *Réponse.* La fonction $x(t)$ est continue partout sauf aux points pour lesquels tous les signes $t_{n_1}, t_{n_2}, \dots$, à partir d'un quelconque, sont les neufs.

16. *Indication.* Si l'arc $\Delta_0 \subset \Gamma$ ne contient pas de points de l'ensemble M , alors les arcs $\Delta_1 = \Delta_0 - 1$, $\Delta_2 = \Delta_0 - 2$, etc., n'en contiennent pas non plus. Les arcs $\Delta_0, \Delta_1, \Delta_2, \dots$ de même longueur sur la circonférence Γ doivent forcément se couper. Si, par exemple, l'arc Δ_k coupe l'arc Δ_{k+m} , alors la réunion des arcs $\Delta_k, \Delta_{k+m}, \Delta_{k+2m}, \dots$ recouvre toute la circonférence Γ .

17. *Indication.* Soit q un point quelconque de l'ensemble $M - P$, et soit $p_1, p_2, \dots$ une suite de points de P qui converge vers q . Montrer, en partant de la continuité uniforme de la fonction $f(x)$ sur P , que la suite $f(p_n)$ a une limite qui ne dépend pas du choix de la suite $p_1, p_2, \dots$ convergente vers le point q . En posant $f(q) = \lim_{n \rightarrow \infty} f(p_n)$ vérifier la continuité de cette fonction sur tout le M .

18. *Indication.* Les intervalles $(f(\alpha-0), f(\alpha+0))$ ne se coupent pas lorsque les α sont différents.

Chapitre 6

1. *Indication.* Appliquer 6.16. en posant $v_n = \frac{1}{n \ln n}$. Utiliser 5.59(6), (11) et (13).

2. *Indication.* Se servir des critères de § 6.1 et de l'exercice 1.

3. *Indication.* Quels que soient n et $m < n$, on a $(n-m)a_n \leq a_{m+1} + \dots$

$$\dots + a_n \leq r_m = \sum_{m+1}^{\infty} a_n, \text{ d'où } na_n \leq \frac{n}{n-m} r_m.$$

4. *L'une des résolutions.* Supposons qu'on ait trouvé, pour tout $n = 1, 2, \dots$, un numéro $N = N(n)$ tel que $\sum_{m=N}^{\infty} a_m < \frac{1}{n^3}$; poser $b_m = n$ pour $N(n) \leq m < N(n+1)$.

5. *Indication.* Appliquer l'inégalité $k(n-k) \leq n^2$ ($k = 1, 2, \dots, n-1$).

6. *Indication.* $s_{2n} \leq s \leq s_{2n+1} = s_{2n} + a_{2n+1}$.

7. *Indication.* Appliquer le critère de d'Alembert 6.14. a.

8. *Indication.* L'existence de la limite résulte de ce que la fonction $f(t)$ est non décroissante et bornée; de plus, $L = \lim_{t \rightarrow t} f(t) \leq \sum_0^{\infty} a_n$. D'autre part, pour tout $\varepsilon > 0$, on peut trouver un N tel que $\sum_0^N a_n > \sum_0^{\infty} a_n - \varepsilon$ et puis

un t tel que $\sum_0^N a_n t^n > \sum_0^{\infty} a_n - \varepsilon$.

9. *Indication.* Même idée que dans l'exercice 8.

10. *Indication.* $x_n = P_{n+1}/P_n$.

11. *Indication.* $\log_b p_n = \sum_{k=1}^n \log_b x_k$.

12. *Indication.* Appliquer 5.59 f.

13. *Indication.* Calculer $(1-x)p_n$.

14. *Indication.* Développer $\frac{1}{1 - \frac{1}{p^x}}$ en une progression géométrique.

15. *Indication.* L'identité d'Euler est valable aussi bien pour $x = 1$; dans ce cas, les deux membres de l'égalité sont infinis.

16. *Indication.* Il y a $9^m - 9^{m-1}$ nombres compris entre $10^{m-1} - 1$ et $10^m - 1$ et ne contenant pas le chiffre 9. Par conséquent, la somme cherchée est plus petite que $\frac{9-1}{1} + \frac{9^2-9}{10} + \frac{9^3-9^2}{100} + \dots = 80$.

17. *Indication.* La longueur d'un vecteur est au moins égale à la longueur de sa projection quelconque et au plus à la somme des longueurs de toutes ses projections sur les axes de coordonnées.

18. *Indication.* Tenir compte de la compacité d'une sphère dans E_n .

19. *Indication.* Appliquer le théorème 6.32.a (6.46).

20. *Indication.* Soit $V_1 \supset V_2 \supset \dots$ une suite des angles solides contenant chacun le vecteur q et tels que $\sup_{x \in V_m, |x|=1} \sin(x, e) = \sigma_m \searrow 0$ et $\sum_1^{\infty} \sigma_m < \infty$.

Choisir successivement les termes de la série (1) contenus dans $V_1, V_2, \dots$ et dont la somme des longueurs est comprise entre 1 et 2.

21. *Indication.* Pour $n = 1$, la proposition se confond avec le théorème de Hilemann 6.37. Soient $f \in E_n$ un vecteur quelconque et E_{n-1} son complément orthogonal. Par récurrence, il existe une permutation de la série (1) dont la somme des projections sur E_{n-1} est égale à un vecteur donné dans E_{n-1} . Ensuite, il faut choisir une partie de la série dont les composantes suivant f forment une série convergente vers $+\infty$ et leurs compléments orthogonaux convergent absolument (exercice 20). La partie restante de la série a les composantes suivant f de somme $-\infty$. En les permutant obtenir la projection voulue sur le vecteur f . La permutation ne change pas la somme des projections orthogonales (exercice 19).

22. *Indication.* Appliquer le résultat de l'exercice 21 au complément orthogonal du sous-espace A .

Chapitre 7

1. *Indication.* $f'(x) \equiv 0$.

2. *Indication.* Appliquer le théorème de Lagrange 7.44.

3. *Indication.* Appliquer le théorème de Lagrange 7.44.

4. *Indication.* On a $y'(0) = 0$, mais $y'(x)$ ne tend pas vers 0 pour $x \rightarrow 0$.

5. *Indication.* Il suffit de voir le cas $f'(a) < 0 < f'(b)$ et de démontrer l'existence d'un point $c \in (a, b)$ avec $f'(c) = 0$. Le point en lequel $f(x)$ atteint sa valeur minimale est le point cherché.

6. *Indication.* Démontrer par l'absurde, à l'aide du théorème 7.44.

7. *Indication.* Si l'inégalité cherchée est juste pour deux points $\lambda = (\lambda_1, \dots, \lambda_p) \in R_p$ et $\mu = (\mu_1, \dots, \mu_p) \in R_p$, elle l'est pour tout le segment qui les joint dans R_p . Elle est évidente pour les points de la forme $\lambda^i = (0, \dots, 1, \dots, 0)$; passer maintenant à n'importe quel point $\lambda = (\lambda_1, \dots, \lambda_p)$.

8. *Indication.* L'inégalité

$$\frac{f(x) - f(\alpha)}{x - \alpha} \leq \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \quad (\alpha < x < \beta)$$

est équivalente à la définition d'une fonction convexe dans l'exercice 7.

9. *Indication.* En passant à la limite dans l'inégalité

$$\frac{f(x) - f(\alpha)}{x - \alpha} \leq \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leq \frac{f(\beta) - f(x)}{\beta - x} \quad (\alpha < x < \beta)$$

pour $\beta \searrow x$ on aboutit à l'existence de $f'_+(x)$ et à l'égalité $f(x+0) = f(x)$. Si l'on pose $\alpha \nearrow x$, on obtient l'existence de $f'_-(x)$ et la relation $f(x-0) = f(x)$.

10. *Indication.* C'est une conséquence de l'inégalité de l'exercice 9.

11. *Indication.* Les intervalles $(f'_-(x), f'_+(x))$, construits pour tout couple de points en lesquels $f'(x)$ n'existe pas ne se coupent pas.

12. *Indication.* Il résulte de l'hypothèse que les seuls points communs à la courbe $y = f(x)$ et à sa corde quelconque sont les extrémités de la corde. La courbe est donc soit au-dessus, soit au-dessous de la corde. Passer continûment de la corde dont les extrémités ont a, b pour abscisses à n'importe quelle corde.

13. *Indication.* Si la courbe $y = f(x)$ possède, pour $x > x_0$, un point $(x_1, f(x_1))$ au-dessous de la demi-tangente, alors la courbe passe au-dessous de la corde joignant les points d'abscisses x_0 et x_1 . Appliquer 7.22 (2).

14. *Indication.* Appliquer la règle de L'Hospital aux limites dans l'exercice 10 du chapitre 4.

15. *Indication.* Appliquer la formule de Lagrange au rapport $\frac{f(a+h) - f(a)}{h}$.

16. *Indication.* Le coefficient angulaire de la tangente en $x_n = \frac{b}{n}$ ne dépasse pas celui de la corde menée par les points x_{n+1} et x_n .

17. *Indication.* En fixant m pour un x_0 donné choisissons l'accroissement $h = \pm \frac{1}{4m}$. Alors les accroissements de tous les $\varphi_m(x)$ à partir du m -ième sont nuls.

La fonction $\varphi_{m-1}(x)$ a les intervalles de longueur $\frac{2}{4m}$ sans points anguleux; celui d'eux qui contient le point x_0 contient aussi l'un des intervalles $(x_0, x_0 + \frac{1}{4m})$, $(x_0, x_0 - \frac{1}{4m})$. Sur cet intervalle, toutes les fonctions précédentes $\varphi_k(x)$, $k < m-1$ n'ont pas de points anguleux non plus, leurs accroissements sont égaux en module à celui de la variable indépendante. Définitivement, on a

$$\frac{\Delta f(x)}{\Delta x} = \sum_{k=0}^{m-1} \frac{\Delta \varphi_k(x)}{\Delta x} = \begin{cases} \text{un nombre pair pour } m \text{ pair,} \\ \text{un nombre impair pour } m \text{ impair.} \end{cases}$$

Donc, $\frac{\Delta f(x)}{\Delta x}$ n'a pas de limite pour $\Delta x \rightarrow 0$.

Chapitre 8

1. *Indication.* Toutes les dérivées de la fonction e^{-1/x^2} pour $x = 0$ sont nulles; pour la démonstration, utiliser l'exercice 15 du chapitre 7.

2. *Indication.* Trouver et estimer le terme principal dans la formule de Leibniz.

3. *Indication.* Appliquer un nombre convenable de fois le théorème de Rolle.

4. *Indication.* Considérer un point x_0 en lequel $|f'(x_0)| > M_1 - \varepsilon$, tenir compte du fait que $f(x) - f(x_0) \geq f'(x_0)(x - x_0) - M_2 \frac{(x - x_0)^2}{2}$.

5. *Réponse.*

$$f^{(n)}(x) = \lim_{h \rightarrow 0} \frac{f(x) - n f(x+h) + \frac{n(n-1)}{1 \cdot 2} f(x+2h) - \dots + (-1)^n f(x+nh)}{h^n}.$$

6. *Indication.* Si $f''(x_0) < 0$, la courbe $y = f(x)$ est au-dessous de sa tangente au voisinage du point $x = x_0$ (8.32). Voir l'exercice 13 du chapitre 7.

7. *Indication.* L'existence d'une corde dont au moins une partie est au-dessous de la courbe conduit à celle d'un point $(x_0, f(x_0))$ au voisinage duquel la courbe est au-dessous de sa tangente; $f''(x) > 0$ exclut cette possibilité.

8. *Indication.* Appliquer les formules 8.73 (5); minorer $\sin z$ en module en tenant compte de ce que $\operatorname{ch} y > \operatorname{sh} y$.

9. *Indication.* Pour $|y| > \varepsilon$ se servir de l'estimation résultant de l'exercice 8. Pour $|y| < \varepsilon$ tenir compte de $|\cos x| > 1 - \delta$.

10. *Réponse.* (a) $\omega < 10^{-7}$; (b) $h \leq 1,03$; (c) $n \geq 10$. En particulier, l'erreur commise en remplaçant le nombre e par la somme $1 + \frac{1}{2!} + \dots + \frac{1}{10!}$ ne dépasse pas 10^{-7} .

11. *Réponse.*

	$f''(x_0) > 0$	$f''(x_0) < 0$
n pair		
n impair		

12. *Réponse.* Non. On peut en donner un exemple en se servant de l'exercice 2.

13. *Indication.* Appliquer l'inégalité

$$\frac{1}{(n+k)!} < \frac{1}{n!n^k} \quad (k=1, 2, \dots).$$

14. *Indication.* Si e est rationnel, alors $n!e$ est un nombre entier pour un certain n .

15. *Indication.* Appliquer les formules 8.37 (2') et (5').

16. *Indication.* Dans un voisinage de 0, l'inégalité cherchée est bien juste. Si elle ne l'est pas dans tout l'intervalle $[-2\pi, 2\pi]$, alors il existe un point $x_0 \neq 0$ en lequel l'inégalité devient égalité. Appliquer le théorème de Lagrange 7.44 à l'intervalle $[0, x_0]$.

17. *Indication.* Poser $x = \pi/2$.

Chapitre 9

1. *Indication.* Le coefficient angulaire de la composante linéaire vaut

$$\frac{1}{T} \int_0^T f(\xi) d\xi,$$

où T est la période de la fonction $f(x)$.

2. *Indication.* Considérer le second membre comme fonction (continue) de ξ , en supposant d'abord que $f(x)$ conserve son signe.

3. *Indication.* Si, par exemple, $\varphi(x)$ décroît, alors $\varphi(x) - \varphi(a+0)$ décroît également et est non négative.

4. *Indication.* Se fait par récurrence sur n . Une autre méthode: remplacer b par x , dériver par rapport à x et voir si les réponses se confondent. Revenir à la formule initiale en intégrant, remarquer que, pour $x = a$, les deux membres de l'égalité coïncident.

5. *Indication.* Généraliser de manière adéquate la démonstration du théorème 9.14.e.

6. *Indication.* Appliquer le critère de Riemann (exercice 5).

7. *Indication.* Vu que

$$||f(x')| - |f(x'')|| \leq |f(x') - f(x'')|,$$

appliquer le critère de Riemann.

8. *Indication.* Appliquer le critère de Riemann.

9. *Indication.* Dans le critère de Du Bois-Reymond (exercice 8), poser $e = 1, \frac{1}{2}, \frac{1}{3}, \dots$ et remplacer δ par $\frac{\delta}{2}, \frac{\delta}{4}, \frac{\delta}{8}, \dots$ respectivement.

10. *Indication.* Appliquer le critère de Lebesgue.

11. *Indication.* Dans l'inégalité

$$\left[\frac{(2n)!}{(2n-1)!} \right]^2 \frac{1}{2n+1} < \frac{\pi}{2} < \left[\frac{(2n)!}{(2n-1)!} \right]^2 \frac{1}{2n},$$

le rapport des membres extrêmes tend vers 1.

12. *Indication.* D'après un $\varepsilon > 0$ donné, trouvons un $\delta(\varepsilon)$ à partir de la condition de la continuité uniforme de la fonction $f(x)$ sur l'intervalle $[A, B]$. Pour $d(\Pi) < \delta$, nous avons

$$f(\xi_j)(x_{j+1} - x_j) = \int_{x_j}^{x_{j+1}} f(x) dx + e_j |x_{j+1} - x_j|,$$

où $|e_j| < \varepsilon$.

13. *Indication.* Si $\sum_{j=0}^{n-1} |x_{j+1} - x_j|$ n'est pas borné, les erreurs commises

en remplaçant $f(\xi_j)(x_{j+1} - x_j)$ par l'intégrale correspondante (2) peuvent augmenter indéfiniment en sommant sur j . Si $f(x)$ a au moins un point de discontinuité, l'erreur n'est pas petite.

14. *Indication.* En se servant de 1)-4), vérifier (3) pour toute fonction en escalier $f(x)$. En appliquant 5) et en approchant la fonction donnée continue par morceaux par une fonction en escalier, aboutir à (3).

15. *Indication.* Substituer $t = 1/\sqrt{3}$ dans 9.74 (4).

Chapitre 10

1. *Indication.* Soit $z=1$ un point d'analyticité de la fonction $f(z)$. On a

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(1)}{k!} (z-1)^k \quad (|z-1| < \rho),$$
 où, d'après l'exercice 8 du chapitre 6 :

$$f(1) = \lim_{z \rightarrow 1} f(z) = \sum_{k=0}^{\infty} a_k, \quad \dots, \quad f^{(n)}(1) = \sum_{k=0}^{\infty} k(k-1) \dots (k-n+1) a_k.$$

D'après la conséquence du théorème d'Abel (6.67) et l'exercice 15 du chapitre 7, on a

$$f(e^{i\theta}) = \sum_{k=0}^{\infty} a_k e^{ikh\theta}, \quad \dots, \quad f^{(n)}(e^{i\theta}) = \sum_{k=0}^{\infty} a_k k(k-1) \dots (k-n+1) e^{ikh\theta},$$

d'où la convergence de la série

$$\sum \frac{f^{(n)}(e^{i\theta})}{n!} (z - e^{i\theta})^n$$

pour $|z - e^{i\theta}| < \rho$.

2. *Indication.* Partir de l'égalité

$$\sum_{n=0}^{\infty} a_n z^n - \frac{b}{z - z_0} = \sum_{n=0}^{\infty} b_n z^n,$$

où $\lim_{n \rightarrow \infty} \sqrt[n]{|b_n|} < 1$.

3. *Indication.* Considérer l'application $w = f(z)$ au voisinage du point z_0 .

4. *Indication.* Appliquer le principe du maximum à la fonction $\frac{f(z)}{z}$.

5. *Indication.* Il faut démontrer la relation $(r_1 < r_2 < r_3)$

$$\ln M(r_2) \leq \frac{\ln r_2 - \ln r_1}{\ln r_0 - \ln r_1} \ln M(r_3) + \frac{\ln r_3 - \ln r_2}{\ln r_2 - \ln r_1} \ln M(r_1).$$

Considérer la fonction $z^\alpha f(z)$, où α est un nombre tel que $r_1^\alpha M(r_1) = r_3^\alpha M(r_3)$.

6. *Indication.* Appliquer la méthode de 10.16.

7. *Indication.* Vérifier pour chacune des fonctions $ax, s + b, \frac{1}{z}$, utiliser 10.55.

8. *Indication.* Estimer les expressions a_n dans 10.38 (2).

9. *Indication.* Calculer l'accroissement de l'argument en contournant chaque zéro et chaque pôle le long de la petite circonférence.

10. *Indication.* Se servir de 10.44.

11. *Indication.* Soit $g(z) = \sum_{m=0}^{\infty} \frac{g^{(m)}(a)}{m!} (z-a)^m$; alors

$$|g^{(m)}(z) + \dots + g^{(m+n)}(z)| = \left| \sum_{k=0}^{\infty} (g^{(m)}(a) + \dots + g^{(m+n+k)}(a)) \frac{(z-a)^k}{k!} \right|.$$

12. *Réponses.* a) Tout feuillet du la surface riemannienne de la fonction $w(z)$ est le plan coupé le long d'une ligne joignant les points $+1$ et -1 (par exemple, le long du segment de l'axe réel $[-1, 1]$ ou de son complémentaire par rapport à l'axe); la surface riemannienne est à deux feuillets, qui sont collés suivant les bords de la coupure: le bord inférieur du premier feuillet avec le bord supérieur du deuxième; les points de branchement sont -1 et $+1$ du plan des w . Dans le plan des z , les domaines où la fonction est injective sont: l'intérieur du cercle unité et l'extérieur de ce cercle. b) Dans le plan des z , les domaines, où la fonction est injective sont les bandes $k\pi < z < (k+1)\pi$ ($k=0, \pm 1, \pm 2, \dots$); la fonction inverse $z = \operatorname{Arccos} w$ est à une infinité de déterminations; pour construire la surface riemannienne, il faut couper chaque feuillet du plan des w le long de l'axe réel de $-\infty$ à -1 et de $+1$ à $+\infty$. Le bord supérieur de la moitié gauche de la coupure du k -ième feuillet est collé au bord inférieur de la moitié gauche de la coupure du $(k+1)$ -ième feuillet, le bord supérieur de la moitié droite de la coupure du k -ième feuillet au bord inférieur de la moitié droite du $(k+1)$ -ième feuillet.

13. *Indication.* Se servir de l'expression de $\cos z$ par l'exponentielle.

14. *Indication.* Considérer la fonction $c_1 f_1(z) + \dots + c_n f_n(z)$, où les nombres c_k sont tels que $|c_k| = 1$ et que, au point z_0 de maximum de $|f_1(z)| + \dots + |f_n(z)|$, on a

$$c_1 f_1(z_0) + \dots + c_n f_n(z_0) = |f_1(z_0)| + \dots + |f_n(z_0)|.$$

15. *Indication.* Appliquer 10.49.

16. *Indication.* Considérer $\frac{1}{P(z)}$.

17. *Indication.* $\frac{P'(z)}{P(z)} = \frac{n}{z} + \frac{g(z)}{z^2}$, où $g(z)$ est bornée pour $z \rightarrow \infty$. Intégrer cette relation le long d'une circonférence de rayon assez grand.

18. *Indication.* Si $P(z)$ ne s'annule pas, alors $\frac{1}{P(z)}$ atteint son maximum en un point fini.

19. *Indication.*

$$\frac{A}{2\pi i} \int_{\Gamma_n} \frac{f(\xi) d\xi}{\xi - z} = f(z) + \sum_k \frac{b_k}{z_k - z},$$

où la sommation s'étend sur tous les pôles de $f(z)$ à l'intérieur du contour Γ_n . D'autre part;

$$\begin{aligned} \frac{1}{2\pi i} \int_{\Gamma_n} \frac{f(\xi) d\xi}{\xi - z} &= \frac{1}{2\pi i} \int_{\Gamma_n} \frac{f(\xi) d\xi}{\xi} + \frac{z}{2\pi i} \int_{\Gamma_n} \frac{f(\xi) d\xi}{\xi(\xi - z)} = \\ &= f(0) + \sum_k \frac{b_k}{z_k} + \frac{z}{2\pi i} \int_{\Gamma_n} \frac{f(\xi) d\xi}{\xi(\xi - z)}; \end{aligned}$$

la dernière intégrale tend vers 0 pour $n \rightarrow \infty$.

20. *Indication.* En ce qui concerne les produits infinis, voir l'exercice 10 du chapitre 6. Intégrer la décomposition en éléments simples donnée dans l'exercice 19 et s'y débarrasser des logarithmes.

21. *Indication.* Se servir de l'exercice 9 du chapitre 8. Pour la première décomposition de l'exercice 20, poser $g(z) = \frac{\sin z}{z}$.

22. *Indication.* Considérer la fonction

$$f(z) e^{-ez^\lambda}, \quad \lambda < \frac{\pi}{2\alpha},$$

dans le secteur $|\arg z| \leq \alpha$, $|z| \leq r$ et appliquer le principe du maximum.

23. *Indication.* Appliquer l'exercice 22 à la fonction $f(z) e^{i\gamma z}$ ($\gamma > 0$) dans chaque quadrant du demi-plan supérieur et déduire l'estimation

$$|f(z) e^{i\gamma z}| \leq \max\{c, m_\gamma\}, \quad m_\gamma = \max_r a_r e^{(\gamma - \gamma_r)r}.$$

Si $m_\gamma > c$, alors la fonction $|f(z)|$ atteint son maximum en un point de l'axe imaginaire, ce qui est impossible pour $f(z) \neq \text{const}$ (exercice 3). Donc, $m_\gamma \leq c$, $|f(z)| \leq c |e^{-i\gamma z}|$; passer à la limite pour $\gamma \rightarrow 0$.

24. *Indication.* Appliquer l'exercice 23 et le théorème de Liouville.

Chapitre 11

1. *Indication.* Démontrer la formule pour le triangle intérieur où la fonction à intégrer est continue, puis passer à la limite.

2. *Indication.*

$$\begin{aligned} \int_a^b \frac{f(ax) - f(bx)}{x} dx &= \int_a^b \frac{f(ax)}{x} dx - \int_a^b \frac{f(bx)}{x} dx = \int_{a\alpha}^{a\beta} \frac{f(x)}{x} dx - \\ &- \int_{b\alpha}^{b\beta} \frac{f(x)}{x} dx = \int_{a\alpha}^{b\alpha} \frac{f(x)}{x} dx - \int_{a\beta}^{b\beta} \frac{f(x)}{x} dx = f(\xi) \ln \frac{b}{a} - f(\eta) \ln \frac{b}{a}, \\ a\alpha &< \xi < b\alpha, \quad a\beta < \eta < b\beta. \end{aligned}$$

3. *Réponse.* $I(p) = \frac{1}{2} \sqrt{\frac{\pi}{p}}$ avec une valeur convenable de $\sqrt{p}$. Choisir pour contour celui du secteur formé d'un segment de l'axe des x , d'un segment de la demi-droite $\varphi = \varphi_0$ et de l'arc d'une circonférence centrée à l'origine des coordonnées.

4. *Indication.* Se servir de l'exercice 3.

Réponse. $F_1 = F_2 = \frac{1}{2} \sqrt{\frac{\pi}{2}}.$

5. *Réponse.*

$$I = 2\pi i \sum \text{Res} \left[\frac{P(x)}{Q(x)} \frac{\sin x}{x} \right],$$

où la somme est étendue sur toutes les racines du polynôme $Q(x)$ dans le demi-plan supérieur.

6. *Réponse.* L'intégrale avec la limite infinie n'est pas uniformément convergente par rapport au paramètre ν .

7. *Indication.* Comparer les valeurs de z^{-1} sur l'intervalle $(z, +\infty)$ parcouru dans les deux sens.

Bibliographie

- [1] А. Г. Курош, Лекции по общей алгебре, гл. 5, § 8, стр. 277 и далее, «Наука», 1962.
- [2] Н. Н. Лузин, Функция. Собрание сочинений, т. 3, изд. АН СССР, М., 1959, стр. 319-344.
- [3] А. И. Маркушевич, Краткий курс теории аналитических функций, Гостехиздат, 1957.
- [4] П. С. Новиков, Элементы математической логики, Физматгиз, 1959.
- [5] И. И. Привалов, Теория функций комплексного переменного, М., 1966.
- [6] В. А. Успенский, Лекции о вычислимых функциях, М., 1960, § 12.
- [7] Успехи математических наук 17, вып. 3 (105) (1962), стр. 163-168.
- [8] Г. М. Фихтенгольд, Курс дифференциального и интегрального исчисления, т. 2, п. 319, стр. 146, Гостехиздат, 1966; т. 2, гл. VIII, п. 262, Гостехиздат, 1948; т. 2, «Наука», 1967.
- [9] А. Я. Хинчин, Краткий курс математического анализа, М., 1953, гл. 27.
- [10] Н. Г. Чеботарев, УМН 2, № 2 (1947).
- [11] Н. Г. Чеботарев, Собрание сочинений, т. 1, изд-во АН СССР, 1950.
- [12] Г. Е. Шилов, Математический анализ, III.17-III.24., изд. МГУ, 1967, стр. 5 и далее.
- [13] Г. Е. Шилов, Математический анализ. Конечномерные линейные пространства, «Наука», 1969, 3.12; 4.76.
- [14] Г. Е. Шилов и Б. Л. Гуревич, Интеграл, мера и производная. Общая часть, «Наука», М., 1967, гл. 2, 4.
- [15] N. Bourbaki, Théorie des ensembles, chap. IV, Hermann, Paris.
- [16] N. Bourbaki, Eléments de mathématique, livre III. Topologie générale, Hermann, Paris.
- [17] J. Dieudonné, Elements d'analyse, tome I, ch. 4, Gauthier-Villars, Paris 1968.
- [18] G. H. Hardy, The Integration of Functions of a Single Variable, Cambridge University Press C. F. Clay manage. London. Fetter Lane, E. C. 1916.
- [19] F. Hausdorff, Mengenlehre, Berlin (de Gruyter) 1927.
- [20] A. Kurosh, Cours d'algèbre supérieure, Editions de Moscou, 1971, chap. 5.
- [21] W. Rudin, Principles of Mathematical Analysis, N.Y., McGraw-Hill, 1953.
- [22] G. Springer, Introduction to Riemann surfaces, 1957.

* * *

V. Smirnov, Cours de mathématiques supérieures en trois tomes, Editions de Moscou.

M. Lavrentiev et B. Chabat, Méthodes de la théorie des fonctions d'une variable complexe, Editions de Moscou.

Index

- Abel* 101, 176
Aire
 du cercle 299
 d'une ellipse 301
 d'une figure plane 269, 333
 — — en coordonnées polaires 303
 d'une sphère 314
 d'une surface de révolution 312
 d'un trapèze curviligne 269, 297-298
Angle
 de deux vecteurs 159
 polaire 157
Appartenance asymptotique 109
Application 163
 conforme 377
 continue 163
Archimède 11, 24, 208, 298, 299, 339
Argand 59
Argument
 d'une fonction 56
 d'un nombre complexe 158
Aryabhata 176
Automorphisme
 de l'espace n -dimensionnel 47
 — —, identique 48
 de la structure 41

Barrow 232
Base 47
Bernoulli Jacques 232,
Bernoulli Jean 232, 388
Bolzano 31, 99, 175, 176, 208, 232
Bombelli 59
Borel 100
Berne
 inférieure 21
 supérieure 25
Boule 62
 fermée 62
 ouverte 62
Bourbaki 59, 100, 232
Bürgi 176

Cantor 31, 38, 40, 58, 100, 176
Cartan H. 127
Casorati 388
Caténoïde 314

Cauchy 31, 99, 101, 127, 128, 175, 176, 208, 232, 255, 339, 388, 389, 436
Cayley 59
Cellule 270
 mesure 270
Chatounovski 127
Chemin 330
Clairaut 339
Coefficients de Laurent 372
Cohen P. 59
Compact 92
 chargé 271
Complémentaire d'un ensemble 34
Complétion 89
Conditions de Cauchy-Riemann 345
Constante d'Euler 395
Continu 40
Continuité
 à gauche et à droite 136
 uniforme 132
Contour fermé 333
Convergence uniforme 171
 à l'intérieur d'un domaine 360
Convolution 420, 422
Coordonnées
 polaires 157
 — dans l'espace 162
 d'un vecteur 45, 47
Corps
 des nombres complexes 53
 numérique 14
Correspondance biunivoque 36
Coupe 56
Courbe lisse par morceaux 305
Critère
 d'Abel-Dirichlet de convergence
 uniforme des intégrales 423
 — pour les intégrales impropres 388
 — pour les séries 194
 de Cauchy 180
 — de convergence uniforme 173, 418
 — — — de l'intégrale impropre 391
 — — — d'une série numérique 178, 180
 — — — d'une suite numérique 84, 121

Critère

- de Cauchy pour la limite d'une fonction vectorielle 126
- — suivant une direction 108
- pour une série vectorielle 193
- de comparaison 179, 392
- de d'Alembert 179
- de Du Bois-Reymond 338
- de Hausdorff 95
- intégral de convergence 393
- de Lebesgue 338
- de Leibniz 185
- pour les intégrales 396
- de Raabe 183
- de Riemann 337
- de Weierstrass 201

D'Alembert 177, 388

Darboux 339

Décomposition

- d'une fonction entière en éléments simples 387
- rationnelle en éléments simples 169

Dedekind 31, 176

Demi-tangente

- à droite 219
- à gauche 219

Dépendance linéaire des vecteurs 46

Dérivation

- d'une intégrale par rapport au paramètre 324
- impropre par rapport au paramètre 416
- d'une suite de fonctions 319, 323

Dérivée 211

- à droite 218
- d'une fonction composée 214
- inverse 215
- puissance 216
- des fonctions trigonométriques 217
- à gauche 218
- du logarithme 215
- d'ordre n 233
- partielle 344
- par rapport à un ensemble 340
- seconde 233

Descartes 59

Diamètre 61

Différentielle

- d'une fonction 222, 344
- composée 222

Différentielles d'ordre supérieur 243

Direction 101

Dirichlet 59, 436

Distance 60

- de deux sous-ensembles 99
- d'un point à un ensemble 99

Domaine 64

- connexe 341
- simplement connexe 354

Droite numérique 21

Du Bois-Reymond 339

Engels 209

Ensemble 11

- ayant la puissance du continu 41
- borné 23
- dans un espace métrique 61
- dénombrable 37
- fermé 79
- fini 11
- infini 11
- jordanien 272
- des nombres réels 14
- majoré 14
- minoré 21
- non borné 62
- non dénombrable 40
- ouvert 65
- partout dense 81
- totalement ordonné 41
- vide 11

Ensembles

- équipotents 36
- géométriquement égaux 64

Equation de Laplace 347

Espace

- compact 92
- complet 84
- euclidien n -dimensionnel 49
- localement compact 92
- métrique 80
- précompact 94
- réel n -dimensionnel 45, 49

Euclide 31

Eudoxe 31

Euler 59, 178, 255, 388, 436

Exemple de Van der Waerden 232

Exponentielle 143, 380

— dans le domaine complexe 247

Exposant de la puissance 17

Extrémum local 221

Fermat 211, 232

Fermeture 81

Fluente 339

Fluxion 256, 339

Fonction(s) 53

- analytique 341
- bêta 425
- bornée 110, 124
- en module 110

Fonction(s)

- caractéristique 272
- circulaires 253
- composée 131
- continue
 - à droite 137
 - sur un ensemble 129
 - à gauche 137
 - par morceaux 265
 - en un point 128
- convexe vers le bas 225, 230
- vers le haut 225, 230
- croisante 137
- décroissante 137
- dérivable en un point 211
 - par rapport à un ensemble 341
- de Dirichlet 267
- domaine de définition 56
 - de valeurs 56
- entière 363
- équivalentes suivant une direction 114
- en escalier 266
- gamma 427
 - dans le domaine complexe 434
 - , expression asymptotique 431
 - , formule des compléments 428
- harmonique 347
- homographique 379
- hyperboliques 251
 - inverses 284
- indéfiniment dérivable 244
- infiniment grande 110
 - — négative 110
 - — positive 110
 - par rapport à une autre 113
 - petite 110
 - — par rapport à une autre 113
- injective 379
- intégrable 258
- inverse 139
- lisse 233
 - par morceaux 278
- majorée 110
- minorée 110
- monotone 137
- multivoque 57
- n fois dérivable 233
 - — par rapport à un ensemble 343
- non croisante 137
- suivant une direction 155

Fonction(s)

- non décroissante 137
 - suivant une direction 118
- non négative suivant une direction 110
- périodique 154, 249
- positive suivant une direction 110
- puissance 144, 380, 382
- réelle 366
- de Riemann 174
- trigonométriques 151
 - dans le domaine complexe 247
 - inverses 155
- uniformément continue 132
- univoque 57
- d'une variable réelle 56
- vectorielle 56

Formule

- de Cauchy 357
- de Dirichlet 435
- de Froullani 435
- de Leibniz 233
- de Newton-Leibniz 276, 293, 351
- de Taylor 235
- de Wallis 338

Formules d'Euler 248

Fréchet 100

Frobenius 55, 59

Frontière d'un ensemble 272

Gauss 31, 59, 176, 208, 388

Gentzen 32

Gödel 32

Grandy 177, 208

Graphique d'une fonction 56

Grassmann 59

Gregory 255

Grouhine 58

Groupement de termes d'une série 188

Hadamard 60, 386

Hausdorff 33, 100

Heine 175

Hilbert 9, 32

Homéomorphisme 70

Homothétie 49

Identité d'Euler 207

Inclusion 12

Inégalité

de Cauchy 50

de quadrilatère 61

de triangle 60

de Young 303

- Intégrale(s)**
 curviligne 330
 — de Stieltjes 330
 définie 276
 elliptiques 286
 de Fourier 408
 — spéciales 411
 de Fresnel 435
 impropre absolument convergente 395
 — convergente 390
 — de deuxième espèce 400
 — divergente 390
 — de première espèce 389
 — semi-convergente 396
 — de troisième espèce 403
 indéfinie 276
 de Laplace 411
 le long d'un contour fermé 333
 de Riemann 257
 — sur un compact 270
 —, les limites 258
 — sur un pavé 271
 de Stieltjes 295
 du type Cauchy 354
- Intégration**
 par changement de variable 278, 293, 400
 multiple par parties 337
 par parties 278, 288, 400
 par rapport au paramètre 324, 414
 d'une suite de fonctions 319
- Intersection d'ensembles** 12, 33
- Intervalle** 23, 24
 contigu 80
- Isomorphisme de structures** 41
- Kepler** 288
- Krylov** 249
- Lagrange** 388
- Lebesgue** 100, 339
- Leibnitz** 232, 255, 256, 258, 293, 339, 388, 389
- Lemme**
 sur les boules fermées 87
 de Jordan 410
 sur le recouvrement fini 97
- L'Hospital** 232
- L'Huilier** 232
- Limite**
 inférieure 116, 122
 suivant une direction 101
 — — sur un sous-ensemble 104
 supérieure 118, 122
- Lobatchevski** 13, 59
- Logarithme** 140
 naturel 149
- Longueur**
 d'un arc de courbe 273, 304
 — — comme fonction du paramètre 309
 — d'ellipse 307
 d'une circonférence 307
 d'un vecteur 49
- Louzine** 389
- Lusternik** 95
- Majorant** 15
 intégrable 419
- Maximum local** 221
- Mengoli** 208
- Mesure de Jordan** 272
- Métrique** 60
- Métriques homéomorphes** 71-72
- Minimum local** 221
- Minorant** 21
- Module**
 d'un nombre complexe 158
 — réel 19
- Moore** 127
- Moyenne**
 arithmético-géométrique 127
 intégrale 262
- Multiple entier** 24
- Multiplicité d'une racine** 168
- Némitrovski** 206
- Neper** 176
- Newton** 208, 232, 255, 256, 339, 389
- Nombre(s)**
 complexes 52
 — conjugués 54
 entiers 16
 irrationnels 17
 finis 30
 maximal 19
 minimal 19
 naturels 16
 négatif 19
 non négatif 19
 non positif 19
 opposé 14
 positif 19
 rationnels 17
 réel inverse 14
 transcendants 40
- Nombre e** 122
- Nombre π** 153
- Norme d'un vecteur** 49
- Novikov** 32

Oscillation d'une fonction 134
on un point 338

Palamodov 58

Parabole osculatrice 239

Partie

entière 24
fractionnaire 24
imaginaire 54
réelle 54

Partition

paramètre 256
à points marqués 256
suivante 258

Pavé

fermé 130
ouvert 130

Peano 31

Période d'une fonction 154, 249

Permutation de termes d'une série 189

Picard 376

Point(s)

d'accumulation de l'ensemble 77
de condensation 99
de continuité 128
de convexité dirigée vers le bas 225
— — vers le haut 225
de discontinuité 128
finis 373
à l'infini 373
d'inflexion 225
intérieur 62
isolé 88
singulier éliminable 373
— essentiel 374
— isolé 373
singuliers d'une courbe 309

Pôle d'ordre m 373

Polya 340, 387

Polynôme de Taylor 237

Primitive 275

Principe

d'Archimède 24
de l'argument 386
d'intervalles formés emboîtés de Cantor 28, 30
du maximum 386

Produit(s)

arithmétique de deux ensembles 31
cartésien d'ensembles 56
de convolution 420
d'ensembles 33
infinis 206
de nombres réels 14
scalaire 49

Produit(s)

d'une série par un nombre 186
de séries 186

Prolongement analytique 246, 365

Ptolémée 176

Puissance

arithmétique n -ième d'un ensemble 31

Quotient d'un nombre réel par un autre 16

Racine

d'une fonction analytique 363
de multiplicité k d'une fonction analytique 363
 n -ième 22

Raisonnement par récurrence 18

Rayon

de convergence 202
polaire 158

Réflexivité 36

Regiomontanus 176

Règles de L'Hospital 228

Résidu 367

logarithmique 370

Reste de la formule de Taylor

sous forme intégrale 289
sous forme de Lagrange 237

Réunion d'ensembles 12, 33

Riemann 339

Rolle 232

Rotation 161

Schwarz 386

Seidel 176

Série

absolument convergente 185

bilatérale 197

—, sommation symétrique 198

de fonctions 200

—, convergence uniforme 203

—, somme 200

harmonique 182

hypergéométrique 205

de Laurent 372

—, partie principale 372

—, — régulière 372

numérique 177

—, convergence 177

—, divergence 177

—, somme 177

—, sommes partielles 177

de puissances 201

—, rayon de convergence 204

Série

- semi-convergente 185, 188
- de Taylor 244, 360
- vectorielle 192
- absolument convergente 193

Signes déclinants 27

Signes d'inclusion 12

Smith 127

Sokhotski 388

Somme

- arithmétique de deux ensembles 31
- d'ensembles 33
- intégrale 257
- de nombres réels 14
- de séries 186

Sous-ensemble 12

propre 12

Sphère 62

Stevin 11

Stieltjes 339

Stiefel 176

Stokes 176

Structure mathématique 41

Suite(s) 56

- bornée 121
- de Cauchy 83
- cofinales 89
- convergente 67, 102
- , limite 67
- divergente 67
- de fonctions 170
- majorée 121
- minorée 121
- non croissante 121
- non décroissante 121

Surface de Riemann 383

Symétrie 36, 64

Système

- binnaire 28
- décimal de numération des nombres réels 26
- d'intervalles emboîtés 28
- ternaire 28

Système complet d'axiomes 42

Szegő 340

Tangente 211

Taylor 255

Théorème

- d'Abel 203
- d'Abel-Liouville 286
- de Baire 87
- de Bolzano 163
- de Bolzano-Weierstrass 78
- de Cauchy 223, 355
- de Cauchy-Hadamard 201
- de Dirichlet 190

Théorème

- sur l'existence d'une racine d'un polynôme 164, 387
- de Hausdorff 89
- de Heine 132
- d'Hermite 122
- de Lagrange 224
- de Liouville 363
- de la moyenne 262, 350
- de Phragmén-Lindelöf 388
- des résidus 369
- de Riemann 190
- de Rolle 223
- de Steinitz 208
- de l'unicité d'une fonction analytique 364
- de Weierstrass 132

Torricelli 232

Transformation d'Abel 194

Transitivité 36

Unité 14

asymptotique 114

Valeur

- absolue d'un nombre réel 19
- d'adhérence d'une fonction 117
- inférieure d'une fonction 116
- supérieure d'une fonction 116
- d'une suite de nombres 122
- de points 76
- d'une fonction 56
- moyenne d'une fonction 282

Vallée Poussin 438

Variables

- dépendante 56
- indépendante 56

Vecteur 45

normé 49

unitaire 49

Vecteurs orthogonaux 160

Voisinage d'un point 62

Volume

- d'une boule 318
- d'un ensemble 272

Weierstrass 31, 175, 176, 232, 388

Wessel 59

Zéro 14

- d'une fonction analytique 363
- de multiplicité k d'une fonction analytique 363

Table des matières

AVANT-PROPOS	5
PREMIÈRE PARTIE	
INTRODUCTION À L'ANALYSE	9
CHAPITRE 1. NOMBRES RÉELS	11
§ 1.1. Premières notions sur les ensembles	11
§ 1.2. Axiomes des nombres réels	13
§ 1.3. Conséquences des axiomes d'addition	15
§ 1.4. Conséquences des axiomes de multiplication	16
§ 1.5. Conséquences des axiomes d'ordre	18
§ 1.6. Conséquences de l'axiome de borne supérieure	21
§ 1.7. Principe d'Archimède et ses conséquences	24
§ 1.8. Principe d'intervalles fermés emboîtés de Cantor	28
§ 1.9. Ensemble des nombres réels achevé	30
Exercices	31
Historique	31
CHAPITRE 2. ÉLÉMENTS DE LA THÉORIE DES ENSEMBLES	33
§ 2.1. Opérations sur les ensembles	33
§ 2.2. Equipotence d'ensembles	35
§ 2.3. Ensembles dénombrables	37
§ 2.4. Ensembles ayant la puissance du continu	40
§ 2.5. Notion de structure mathématique. Isomorphisme de structures	41
§ 2.6. Espace n -dimensionnel	45
§ 2.7. Nombres complexes	52
§ 2.8. Notion générale de fonction. Graphique	56
Exercices	57
Historique	58
CHAPITRE 3. ESPACES MÉTRIQUES	60
§ 3.1. Définitions et exemples	60
§ 3.2. Ensembles ouverts	65
§ 3.3. Suites convergentes et homéomorphisme	67
§ 3.4. Valeurs d'adhérence et points d'accumulation	76
§ 3.5. Ensembles fermés	79
§ 3.6. Ensembles partout denses. Fermetures	81
§ 3.7. Espaces complets	83
§ 3.8. Complétion	89
§ 3.9. Compacité	92
Exercices	98
Historique	99
CHAPITRE 4. THÉORIE GÉNÉRALE DES LIMITES	101
§ 4.1. Définition d'une limite	101
§ 4.2. Théorèmes généraux sur les limites	108
§ 4.3. Limites de fonctions numériques	109
§ 4.4. Valeurs d'adhérence d'une fonction	116
§ 4.5. Fonctions non décroissantes suivant une direction	118

§ 4.6. Théorèmes fondamentaux sur des suites numériques	120
§ 4.7. Limites de fonctions vectorielles	124
Exercices	126
Historique	127
CHAPITRE 5. FONCTIONS CONTINUES	128
§ 5.1. Fonctions continues sur un espace métrique	128
§ 5.2. Fonctions numériques continues sur la droite numérique	135
§ 5.3. Fonctions monotones	137
§ 5.4. Logarithme	140
§ 5.5. Exponentielle	143
§ 5.6. Fonctions trigonométriques	151
§ 5.7. Applications des fonctions trigonométriques	157
§ 5.8. Fonctions vectorielles continues d'une variable vectorielle	163
§ 5.9. Suites de fonctions	170
Exercices	174
Historique	175
CHAPITRE 6. SÉRIES	177
§ 6.1. Séries numériques. Séries à termes positifs	177
§ 6.2. Séries à termes réels de signes quelconques	184
§ 6.3. Opérations sur les séries	186
§ 6.4. Séries de vecteurs	192
§ 6.5. Séries de fonctions	200
§ 6.6. Séries de puissances	201
Exercices	205
Historique	208
DEUXIÈME PARTIE	
CALCUL DIFFÉRENTIEL ET INTÉGRAL	209
CHAPITRE 7. DÉRIVÉE	211
§ 7.1. Définition d'une dérivée	211
§ 7.2. Deuxième définition de la dérivée	219
§ 7.3. Différentielle	221
§ 7.4. Théorèmes des accroissements finis	223
§ 7.5. Position d'une courbe par rapport à sa tangente	225
§ 7.6. Règles de L'Hospital	228
Exercices	230
Historique	232
CHAPITRE 8. DÉRIVÉES D'ORDRE SUPÉRIEUR	233
§ 8.1. Définitions et exemples	233
§ 8.2. Formule de Taylor	235
§ 8.3. Analyse du comportement d'une fonction au voisinage d'un point donné	239
§ 8.4. Différentielles d'ordre supérieur	243
§ 8.5. Série de Taylor	244
§ 8.6. Exponentielle et fonctions trigonométriques dans le domaine complexe	247
§ 8.7. Fonctions hyperboliques	251
Exercices	253
Historique	255
CHAPITRE 9. INTÉGRALE DE RIEMANN	250
§ 9.1. Définition de l'intégrale et théorèmes d'existence	250
§ 9.2. A quoi sert l'intégrale?	268

§ 9.3. Intégrale comme fonction de sa limite supérieure	274
§ 9.4. Calcul des intégrales indéfinies	279
§ 9.5. Calcul des intégrales définies	288
§ 9.6. Applications de l'intégrale	297
§ 9.7. Intégration et dérivation d'une suite de fonctions	319
§ 9.8. Intégration et dérivation par rapport au paramètre	324
§ 9.9. Intégrales curvilignes	330
Exercices	337
Historique	339
CHAPITRE 10. FONCTIONS ANALYTIQUES	340
§ 10.1. Définitions et exemples	340
§ 10.2. Intégrales curvilignes de fonctions complexes	347
§ 10.3. Théorème de Cauchy et ses conséquences	354
§ 10.4. Résidus et points singuliers isolés	367
§ 10.5. Applications du plan complexe dans lui-même. Fonctions élémentaires	377
Exercices	385
Historique	388
CHAPITRE 11. INTÉGRALES IMPROPRES	389
§ 11.1. Intégrales impropres de première espèce	389
§ 11.2. Intégrales impropres de deuxième et de troisième espèce	400
§ 11.3. Calcul d'intégrales impropres à l'aide de résidus	404
§ 11.4. Intégrales impropres de fonctions d'un paramètre	413
§ 11.5. Fonctions gamma et bêta d'Euler	424
Exercices	435
Historique	436
INDICATIONS ET RÉPONSES	437
Bibliographie	447
Index	448

A NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur la traduction et la présentation de ce livre, ainsi que toute autre suggestion que vous voudriez formuler.

Ecrire à l'adresse :
Pervi Rijski péréoulouk, 2,
Moscou, I-110, GSP, U.R.S.S.

Imprimé en Union Soviétique